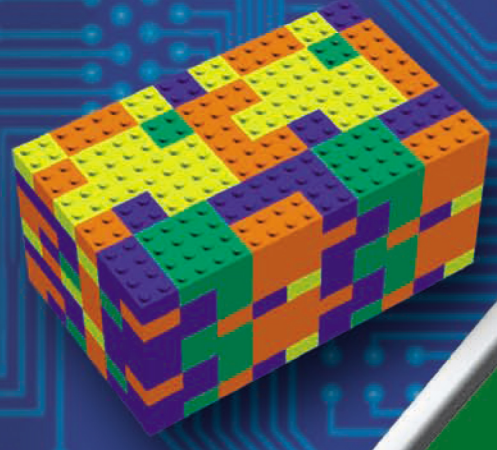
Virtualização

Componente Central do Datacenter

Manoel Veras

- Explica o que é VIRTUALIZAÇÃO de servidores
- Integra o conceito de VIRTUALIZAÇÃO com os conceitos de DATACENTER e CLOUD COMPUTING
- Ensina a fazer um projeto completo de VIRTUALIZAÇÃO
- Orienta o uso das principais ferramentas de apoio a projetos de VIRTUALIZAÇÃO
- Aborda o principal software de mercado, o VMware vSphere, sugerindo também arquiteturas de referência



Prefácio
Marco Américo D. Antonio
Gerente Geral da Diveo do Brasil

BRASPORT

Manoel Veras

Virtualização

Componente Central do Datacenter

Prefácio
Marco Américo D. Antonio
Gerente Geral da Diveo do Brasil

BRASPORT

Copyright© 2011 por Brasport Livros e Multimídia Ltda.
Todos os direitos reservados. Nenhuma parte deste livro poderá ser reproduzida, sob qualquer meio, especialmente em fotocópia (xerox), sem a permissão, por escrito, da Editora.

Impressão: 2011

Editor: Sergio Martins de Oliveira
Diretora: Rosa Maria Oliveira de Queiroz
Gerente de Produção Editorial: Marina dos Anjos Martins de Oliveira
Revisão: Maria Inês Galvão
Editoração Eletrônica: Abreu's System Ltda.
Capa: Paulo Vermelho

Técnica e muita atenção foram empregadas na produção deste livro. Porém, erros de digitação e/ou impressão podem ocorrer. Qualquer dúvida, inclusive de conceito, solicitamos enviar mensagem para **brasport@brasport.com.br**, para que nossa equipe, juntamente com o autor, possa esclarecer. A Brasport e o(s) autor(es) não assumem qualquer responsabilidade por eventuais danos ou perdas a pessoas ou bens, originados do uso deste livro.

BRASPORT Livros e Multimídia Ltda.
Rua Pardal Mallet, 23 – Tijuca
20270-280 Rio de Janeiro-RJ
Tels. Fax: (21) 2568.1415/2568.1507
e-mails: brasport@brasport.com.br
vendas@brasport.com.br
editorial@brasport.com.br

site: **www.brasport.com.br**

Filial
Av. Paulista, 807 – conj. 915
01311-1))0 – São Paulo-SP
Tel. Fax (11): 3287.1752
e-mail: filialsp@brasport.com.br

Dedico esta obra aos meus pais João Hélio (*in memorian*) e Vera Maria.

"Mais cedo ou mais tarde, algum ponto fundamental em seu negócio mudará".

(Andrew S. Grove, EX-CEO Intel).

Agradecimentos

Agradeço a todos que me incentivam a prosseguir. Este livro VIRTUALIZAÇÃO é um desdobramento do livro DATACENTER, mas não menos importante. Estruturá-lo e torná-lo prático sem esquecer dos fundamentos teóricos do tema foi um grande desafio. O ponto certo entre a teoria sobre VIRTUALIZAÇÃO e os aspectos práticos foi o que busquei.

Gostaria de agradecer a todos que contribuíram e me convenceram de que valeria a pena enfrentar esta nova jornada para a obtenção deste objetivo. Sou grato a diversos profissionais e acadêmicos da computação brasileira com quem convivi nestes últimos anos e que me ajudaram a tornar este sonho uma realidade.

Gostaria de agradecer ao Luiz Augusto e ao Geraldo Marcelo pela contribuição no caso vSphere.

Gostaria também de agradecer ao Sergio Martins e à Rosa Queiroz, da Editora Brasport, pelo apoio dado ao projeto desde o início.

Manoel Veras

Nota do Autor

As sugestões feitas neste livro devem ser tratadas como linhas orientadoras para profissionais que buscam atuar nas áreas de conhecimento envolvidas e necessitam de um referencial que vincule a Tecnologia da Informação com a VIRTUALIZAÇÃO de servidores. As áreas envolvidas diretamente neste livro são: Tecnologia da Informação (TI), Infraestrutura de TI, CLOUD COMPUTING (COMPUTAÇÃO DE NUVEM), DATACENTERS e VIRTUALIZAÇÃO, normalmente abordadas com diferentes enfoques na Ciência da Computação, Engenharia da Computação, Sistemas de Informação e até na Administração de Empresas.

Este é um livro com foco na VIRTUALIZAÇÃO de servidores, principal componente do DATACENTER do futuro. A VIRTUALIZAÇÃO acontece em todos os níveis da infraestrutura de TI, mas é no servidor que a vantagem de utilizá-la é mais clara. O que também acontece é que os softwares de VIRTUALIZAÇÃO acabam permitindo que a VIRTUALIZAÇÃO de servidores seja estendida para o armazenamento e para a rede.

Por sua vez, a CLOUD COMPUTING só é efetiva quando possui o(s) DATACENTER(S), seu(s) principal(is) componente(s), provido(s) com recursos de VIRTUALIZAÇÃO. A abordagem utilizada no livro é centrada nos aspectos práticos da VIRTUALIZAÇÃO, no projeto e também no entendimento das principais funcionalidades de softwares de VIRTUALIZAÇÃO disponíveis no mercado.

Diversas publicações sugerem que o mercado de trabalho para profissionais que lidam com os assuntos aqui tratados só tende a crescer. As organizações precisam de profissionais que entendam o papel da infraestrutura de TI, da CLOUD COMPUTING, do DATACENTER e da VIRTUALIZAÇÃO na nova organização. Estes profissionais também precisam ter uma visão clara sobre as tecnologias de VIRTUALIZAÇÃO, sobre os impactos na arquitetura e na infraestrutura a ser demandada pelas aplicações e consequentes cargas de trabalho (WORKLOADS) e sobre as reais vantagens do seu uso para a organização. Por fim, precisam projetar estas novas estruturas e conduzir de maneira efetiva os projetos que estão no centro das decisões de TI.

Prefácio

É com muita alegria que apresento aqui a mais nova obra do Prof. Dr. Manoel Veras, a qual vem para complementar com maior profundidade e especificidade seu livro anterior, que abordava os principais conceitos ligados ao mundo dos DATACENTERS.

Nesta obra, o tema VIRTUALIZAÇÃO é enfocado com grande ênfase e profundidade, cobrindo com detalhes todo o lado técnico mas não esquecendo de discutir os principais impactos práticos no dia a dia dos negócios.

Atualmente, a VIRTUALIZAÇÃO já é aplicada em larga escala nos mais diversos projetos: na consolidação de ambientes, no provimento dos mais diversos serviços na nuvem acessada pela Internet e na criação de nuvens computacionais privadas utilizadas globalmente por grandes corporações.

O que as empresas buscam com a utilização desta tecnologia é a diminuição do consumo de energia elétrica, a redução no consumo de espaço físico, a racionalização dos recursos computacionais existentes com o aumento de performance e a disponibilidade de suas aplicações, sem deixar de lado a preocupação com a segurança dos dados que aí estão sendo armazenados e processados.

Na prática, a VIRTUALIZAÇÃO permite que seja feito muito mais com menos, reduzindo não só os custos operacionais e investimentos necessários mas também os impactos socioambientais gerados pela indústria de TI e dos DATACENTERS.

Como profissional especializado, eu gostaria de ressaltar que esta obra, além de possuir um caráter acadêmico, pode e deve ser utilizada como referência no dia a dia dos negócios no mercado corporativo.

Gostaria também de fazer uma citação pessoal ao Prof. Dr. Manoel Veras pelo seu conhecimento, interesse e visão desta indústria bem como pelo seu lado empreendedor na criação e publicação de um material tão rico e complexo.

Tenho certeza de que todos os seus leitores aproveitarão muito do conteúdo de seu livro e espero que o professor Dr. Manoel Veras continue nos brindando com novas e elucidativas publicações.

Marco Américo D. Antonio
Gerente Geral da Diveo do Brasil

Sumário

Introdução 1
Objetivos 1
Estrutura 2

PARTE I: VISÃO GERAL

1. Infraestrutura de TI e Virtualização 7
1.1. Introdução 7
1.2. Investimentos em Infraestrutura de TI 9
1.3. Maturidade da Infraestrutura de TI 12
1.3.1. Introdução 12
1.3.2. Modelo de Weill e Broadbent 13
1.3.3. Modelos da Microsoft 14
1.4. Gerenciamento de Serviços de TI 17
1.4.1. ITIL 18
1.4.2. ITIL v3 18
1.4.3. TCO 21
1.4.4. ITIL e TCO 22
1.5. Impactos da VIRTUALIZAÇÃO na Infraestrutura de TI 22
1.6. Referências Bibliográficas 23

2. Cloud Computing e Virtualização 25
2.1. Introdução 25
2.2. Conceito 27
2.3. Características Essenciais 29
2.4. Modelos de Serviços 29
2.5. Modelos de Implantação 32
2.6. Benefícios 33
2.7. Desafios 33
2.8. Iniciativas 34
2.8.1. Conceito na Prática: VMware vSphere 34

2.8.2. Conceito na Prática: Windows AZURE 36
2.8.3. Conceito na Prática: Amazon AWS 40
2.9. Riscos e Segurança 42
2.10. Impactos da VIRTUALIZAÇÃO na CLOUD COMPUTING 46
2.11. Referências Bibliográficas 47

3. Tier Datacenter e Virtualização 48
3.1. Introdução 48
3.2. TIER DATACENTER 53
3.2.1. TIER 1 56
3.2.2. TIER 2 56
3.2.3. TIER 3 56
3.2.4. TIER 4 57
3.3. TCO do DATACENTER 58
3.4. Padronização do DATACENTER 59
3.4.1. CHASSI 59
3.4.2. RACK 59
3.4.3. CONTAINER 61
3.4.4. Conceito na Prática: DATACENTER POD da HP 62
3.4.5. Conceito na Prática: SALA COFRE da Aceco 63
3.5. Construção do DATACENTER 64
3.6. Impactos da VIRTUALIZAÇÃO no DATACENTER 65
3.7. Referências Bibliográficas 66

4. Green Datacenter e Virtualização 67
4.1. Introdução 67
4.2. Eficiência do DATACENTER 67
4.3. O Green Grid 71
4.4. Conceito na Prática: APC TradeOff Tools 73
4.5. Requisitos de Energia e Refrigeração para DATACENTERS 77
4.5.1. Carga de TI 77
4.5.2. No-Break 78
4.5.3. Iluminação 78
4.5.4. Carga da Refrigeração 78
4.5.5. Dimensionamento do Sistema Elétrico 78
4.6. Medição do PUE 79
4.7. Impactos da VIRTUALIZAÇÃO no Green DATACENTER 81
4.8. Referências Bibliográficas 81

PARTE II: TECNOLOGIA DE VIRTUALIZAÇÃO

5. Conceitos Centrais 85
5.1. Introdução 85

5.2. Workload e Throughput 86
5.3. Conceito 87
5.4. Máquina Virtual 88
5.5. Histórico 89
5.6. Benefícios 91
5.7. Principais Fornecedores 93
5.8. Desempenho e Benchmarks 94
5.9. Limitações 95
5.10. Utilização 95
5.11. Referências Bibliográficas 96

6. Técnicas 97
6.1. Introdução 97
6.2. Hardware 97
6.3. Sistema Operacional (SO) 98
6.4. Categorias de Virtualização 100
6.5. Máquina Virtual de Sistema ou HYPERVISORS 101
6.5.1. TIPO UM – Baremetal 101
6.5.2. TIPO DOIS 103
6.6. VIRTUALIZAÇÃO Total 104
6.7. PARAVIRTUALIZAÇÃO 106
6.8. VIRTUALIZAÇÃO Assistida por Hardware 108
6.9. Comparação entre as Técnicas de VIRTUALIZAÇÃO 108
6.10. Referências Bibliográficas 109

7. Arquitetura e Infraestrutura 110
7.1. Introdução 110
7.2. Arquitetura Física do DATACENTER 111
7.3. Arquitetura Virtual do DATACENTER 113
7.4. Blades, Servidores e VIRTUALIZAÇÃO 116
7.4.1. Introdução 116
7.4.2. Conceito na Prática: Dell Blade 1000E 120
7.4.3. Processadores 125
7.4.4. Conceito na Prática: Intel Processador XEON 5600 126
7.4.5. Licenciamento 128
7.4.6. Conceito na Prática: Calculadoras do Windows Server 2008 130
7.5. Rede e VIRTUALIZAÇÃO 132
7.5.1. Conceito na Prática: Cisco Nexus 1000V 134
7.6. Storage e VIRTUALIZAÇÃO 136
7.6.1. Introdução 136
7.6.2. Discos e Interfaces 136
7.6.3. Sistema de Armazenamento 137
7.6.4. Protocolo SCSI 137
7.6.5. Características do Storage 138

7.6.6. Tipos de Storage....139
7.6.7. Redes de Storage....140
7.6.8. SAN FC....142
7.6.9. SAN IP....143
7.6.10. Conceito na Prática: Dell Equallogic PS Series....146
7.6.11. NAS....147
7.6.12. SAN FCoE....148
7.6.13. Conceito na Prática: Unified Storage EMC Celerra NS-120....151
7.6.14. VIRTUALIZAÇÃO do Storage....154
7.6.15. Conceito na Prática: EMC VPLEX....156
7.7. Alta Disponibilidade e Recuperação de Desastres e VIRTUALIZAÇÃO....159
7.7.1. Conceitos Importantes....160
7.7.2. Clusters de Servidores....163
7.7.3. Backup e Restore....165
7.7.4. Desduplicação....168
7.7.5. Arquivamento (Archive)....169
7.7.6. Replicação....171
7.7.7. Conceito na Prática: EMC RecoverPoint....174
7.7.8. Replicação com Desduplicação....175
7.8. Desktop e VIRTUALIZAÇÃO....176
7.9. Gerenciamento e VIRTUALIZAÇÃO....177
7.10. Segurança e VIRTUALIZAÇÃO....177
7.11. Conceito na Prática: Cisco DATACENTER 3.0....178
7.11.1. Portfólio....178
7.11.2. Roadmap....178
7.11.3. Componentes....179
7.12. Referências Bibliográficas....182

PARTE III: SOFTWARE DE VIRTUALIZAÇÃO

8. VMware vSphere....185
8.1. Introdução....185
8.2. Benefícios....186
8.3. Utilização....187
8.4. Componentes da Solução....188
8.4.1. Serviços de Infraestrutura....188
8.4.2. Serviços de Aplicativos....191
8.5. Novidades com o vSphere 4.1....193
8.5.1. Introdução....193
8.5.2. Storage I/O Control (SIOC)....193
8.5.3. Network I/O Control (NIOC)....194
8.5.4. vSphere vCenter Server 4.1....195
8.5.5. Configurações Máximas do vSphere 4.1....197

8.6. Edições 201
8.6.1. Características 201
8.6.2. Kits e Edições 201
8.6.3. Add-ons e Produtos Adicionais 202
8.7. ESX e ESXi 203
8.7.1. Introdução 203
8.7.2. Funcionamento 205
8.7.3. Diferenças 206
8.7.4. Recursos 206
8.7.5. Arquitetura 207
8.7.6. Gerenciamento 208
8.7.7. Desempenho e Escalabilidade 210
8.7.8. Disponibilidade 211
8.7.9. Interoperabilidade 211
8.7.10. Segurança 212
8.8. Sistema de Arquivos vStorage VMFS 213
8.8.1. Introdução 213
8.8.2. Funcionamento 214
8.8.3. Recursos 214
8.8.4. RDM 216
8.9. vMotion 216
8.9.1. Introdução 216
8.9.2. Utilização 216
8.9.3. Funcionamento 217
8.9.4. Recursos 218
8.10. vCenter Server 219
8.10.1. Conceito 219
8.10.2. Utilização 219
8.10.3. Funcionamento 220
8.10.4. Características 220
8.10.5. vCenter Server Heartbeat 221
8.10.6. vCenter Site Recovery Manager (SRM) 223
8.11. Funcionalidades Importantes do VMware vSphere 225
8.11.1. VMware High Availability (HA) 225
8.11.2. VMware Distributed Resource Scheduler (DRS) 226
8.11.3. VMware Distributed Power Management (DPM) 226
8.11.4. VMware Fault Tolerance (FT) 227
8.11.5. VMware Data Recovery (VDR) 228
8.11.6. VMware Storage VMotion 231
8.11.7. VMware View 232
8.12. Referências Bibliográficas 234

PARTE IV: PROJETO

9. Metodologia e Ferramentas de Projeto 239
9.1. Introdução 239
9.2. Gestão de Projetos de VIRTUALIZAÇÃO 241
9.2.1. Projeto do DATACENTER 241
9.2.2. Projeto da VIRTUALIZAÇÃO 244
9.3. Projeto 1 – Business Case e Escolha do Fornecedor 245
9.4. Business Case 246
9.4.1. Introdução 246
9.4.2. Situação Atual 247
9.4.3. Desafios 247
9.4.4. Benefícios 247
9.4.5. Análise de TCO/ROI 249
9.5. Escolha do Fornecedor 250
9.5.1. Custo e Desempenho 250
9.5.2. Funcionalidades 254
9.5.3. Base Instalada 260
9.5.4. Suporte 260
9.6. Projeto 2 – Solução de VIRTUALIZAÇÃO 260
9.7. Iniciação 260
9.8. Planejamento 261
9.8.1. Documentação de Requisitos 261
9.8.2. Declaração do Escopo 261
9.8.3. Criação da EAP 261
9.8.4. Atividades em Termos de Tempo, Custo e Qualidade 262
9.8.5. Identificação dos Riscos 262
9.8.6. Análise da Situação Atual 262
9.8.7. Considerações sobre a DMZ 263
9.8.8. Considerações sobre a VIRTUALIZAÇÃO das Principais Aplicações 263
9.8.9. Benchmarking da Indústria de TI 265
9.8.10. Opções para Definir a Infraestrutura 269
9.8.11. Conceito na Prática: Microsoft SQL Server 272
9.8.12. Conceito na Prática: Oracle Database 280
9.8.13. Conceito na Prática: Microsoft Exchange Server 282
9.8.14. Conceito na Prática: ERP SAP 288
9.8.15. Conceito na Prática: Dell Virtualization Advisor Tool 293
9.9. Implementação 294
9.9.1. Introdução 294
9.10. Definição de Arquitetura e Infraestrutura 295
9.10.1. Introdução 295
9.10.2. Arquiteturas Típicas 296
9.10.3. Conceito na Prática: Dell VMware vSphere 298
9.10.4. Conceito na Prática: Cisco EMC VMware vSphere Vblock 302

9.10.5. Vblock 0 304
9.10.6. Conceito na Prática: EMC Ionix 306
9.10.7. Aquisição de Infraestrutura e Serviços 307
9.10.8. Instalação, Configuração, Testes e Validação 307
9.11. Monitoramento e Controle 308
9.11.1. Definição dos Níveis de Serviço e Indicadores 308
9.12. Encerramento 308
9.12.1. Lições Aprendidas 308
9.13. Referências Bibliográficas 308

10. Case vSphere 309
10.1. Introdução 309
10.2. Business Case, Escolha do Fornecedor e Versão do Software 309
10.3. Documentação de Requisitos 309
10.4. Declaração do Escopo 311
10.4.1. Objetivo Geral 311
10.4.2. Objetivos Específicos 311
10.4.3. Pressupostos 311
10.4.4. Demarcações e Regras 312
10.4.5. Identificação de Riscos 312
10.5. Situação Atual 312
10.5.1. Inventário 312
10.5.2. Análise da Demanda das WORKLOADS 313
10.6. Definição da Arquitetura e da Infraestrutura 315
10.6.1. Especificações Gerais do Projeto 315
10.7. Dimensionamento da Infraestrutura 321
10.7.1. Dimensionamento dos Servidores 321
10.7.2. Dimensionamento do Armazenamento 322
10.7.3. Dimensionamento da Alta Disponibilidade 322
10.7.4. Necessidade do Uso de Recursos 323
10.7.5. Desenho do Armazenamento 323
10.7.6. Dimensionamento da DMZ 324
10.7.7. Desenho do Gerenciamento e Monitoramento 324
10.8. Definição dos Níveis de Serviço e Indicadores 325
10.8.1. Diagramas 327
10.9. Referências Bibliográficas 330

Índice Remissivo 331

Introdução

Este livro representa o balanço de anos de experiência nesta área. São tratados aspectos importantes que contribuem para a formação de profissionais na área de INFRAESTRUTURA DE TECNOLOGIA DA INFORMAÇÃO (TI) com foco na VIRTUALIZAÇÃO do DATACENTER. Particularmente, trabalhei como consultor em algumas dezenas de projetos de VIRTUALIZAÇÃO envolvendo o projeto e a definição da infraestrutura para as soluções da VMware e da MICROSOFT.

Qual a linha de base estabelecida para o livro? Partiu-se do genérico, associando a infraestrutura de TI à VIRTUALIZAÇÃO de servidores, indo até o específico, tratando de questões puramente técnicas relacionadas ao projeto e tecnologias de VIRTUALIZAÇÃO. Uma dificuldade natural de um livro com este foco é conseguir sequenciar os assuntos, de forma a fazer com que o leitor avance passo a passo. Procurou-se construir os assuntos na melhor sequência possível, mas eventualmente é preciso chamar um conceito que só será explicado posteriormente. Este aspecto deve ser considerado durante a leitura.

Objetivos

Diversas publicações com foco na área de TI salientam o surgimento de novas profissões oriundas das mudanças ocorridas no formato das organizações e no papel que a TI passa a desempenhar. A organização se torna cada vez mais baseada na informação e a TI permite levar a informação para onde interessa, no tempo que interessa. As aplicações baseadas em TI são a coluna vertebral da nova organização e a infraestrutura é que permite obter o desempenho e a disponibilidade necessários para estas aplicações.

Com a tendência de consolidação da infraestrutura de TI e o surgimento de novas tecnologias de servidores baseadas em plataformas abertas de baixo custo, surge a VIRTUALIZAÇÃO como tema central na construção desta nova infraestrutura. Em um mundo dinâmico onde as organizações e os clientes se reconfiguram a cada momento, é necessário que as aplicações que suportam estes processos atendam rapidamente a estas novas demandas. Uma das formas de tentar atender a novas demandas por processos mais flexíveis é construir aplicações orientadas a serviços. Aplicações orientadas a serviços (*Services Oriented Application* – SOA), por sua vez, reforçam a necessidade de orientar também a infraestrutura para serviços (*Services Oriented Infrastructure – SOI*) para dar suporte a esta nova organização. Mas isto só não basta.

A VIRTUALIZAÇÃO de servidores em arquiteturas abertas e de baixo custo, como a arquitetura x86, vem resolver um problema chave relacionado à construção desta infraestrutura. Com a VIRTUALIZAÇÃO, pode-se adequar o hardware à carga de trabalho originada pela aplicação. Ou seja, se a aplicação demanda muito recurso computacional pode-se alterar a configuração do DATACENTER dinamicamente e fornecer este recurso subtraindo o recurso de aplicações que naquele momento são pouco demandantes. Em outro momento, pode-se novamente restabelecer a configuração original e até expandir o uso do recurso. A VIRTUALIZAÇÃO passa a ser a base para a infraestrutura.

O objetivo principal deste livro é contribuir para a formação de arquitetos de DATACENTERS que possuam como elemento central a VIRTUALIZAÇÃO. Um objetivo específico é fazer com que o livro funcione como um guarda-chuva de informação sobre esta área de conhecimento, integrando diversos assuntos envolvidos, tratando o assunto em pauta com certa profundidade, mas também permitindo que o leitor avance em um tema específico ao consultar o vasto material citado como referência e sites de consulta.

Estrutura

O livro é dividido em quatro partes: Visão Geral, Tecnologia de VIRTUALIZAÇÃO, Software e Projeto. A ideia é que os assuntos tratados nas partes/capítulos tenham certa independência, mesmo que fazendo parte de uma sequência lógica e assim permitindo que o leitor possa ler um único capítulo ou mesmo uma parte específica.

A PARTE I, "VISÃO GERAL", trata de contextualizar a importância da VIRTUALIZAÇÃO para a infraestrutura de TI, sua relação com a CLOUD COMPUTING e com o DATACENTER. Também ilustra a sua importância para a nova organização baseada na informação.

A PARTE II, "TECNOLOGIA", aborda os conceitos centrais e as técnicas sobre VIRTUALIZAÇÃO envolvendo HYPERVISORS e a arquitetura e a infraestrutura necessárias.

A PARTE III, "SOFTWARE ", apresenta o principal software de mercado, suas funcionalidades e arquitetura.

A PARTE IV, "PROJETO", traz uma metodologia de projeto para a VIRTUALIZAÇÃO e ilustra o uso de ferramentas de projeto fornecidas pelos principais fabricantes. Também apresenta um caso real de VIRTUALIZAÇÃO construído com vSphere que utiliza a metodologia de projeto sugerida.

PARTE I: VISÃO GERAL

Capítulo 1: INFRAESTRUTURA DE TI E VIRTUALIZAÇÃO

Capítulo 2: CLOUD COMPUTING E VIRTUALIZAÇÃO

Capítulo 3: TIER DATACENTER E VIRTUALIZAÇÃO

Capítulo 4: GREEN DATACENTER E VIRTUALIZAÇÃO

PARTE II: TECNOLOGIA DE VIRTUALIZAÇÃO

Capítulo 5: CONCEITOS CENTRAIS

Capítulo 6: TÉCNICAS

Capítulo 7: ARQUITETURA E INFRAESTRUTURA

PARTE III: SOFTWARE DE VIRTUALIZAÇÃO

Capítulo 8: VMWARE VSPHERE

PARTE IV: PROJETO

Capítulo 9: METODOLOGIAS E FERRAMENTAS DE PROJETO

Capítulo 10: CASE VSPHERE

Vale salientar que o aspecto prático é sempre considerado e o livro traz diversos exemplos de casos e dicas de implementações reais das tecnologias citadas.

Existem diversos softwares utilizados em projetos de VIRTUALIZAÇÃO, como o MICROSOFT WINDOWS SERVER 2008 (WS08) com HYPER V, CITRIX XEN SERVER e o PARALLELS SERVER, dos quais algumas características são comentadas no texto, mas optou-se neste livro por dar ênfase ao principal software de VIRTUALIZAÇÃO do mercado, o VMware vSphere, cuja documentação existente permitiu detalhar a sua arquitetura e suas principais funcionalidades.

Importante deixar claro que as seções CONCEITO NA PRÁTICA são baseadas em informações fornecidas pelos fabricantes em artigos públicos, folhas de especificação (*spec sheets*) ou em seus sites específicos e não são originadas pelo autor. Optou-se também por não traduzir a palavra HYPERVISOR. APLICATIVOS e APLICAÇÕES são termos utilizados como sinônimos no livro.

A **Figura 0-1** ilustra o sequenciamento do livro.

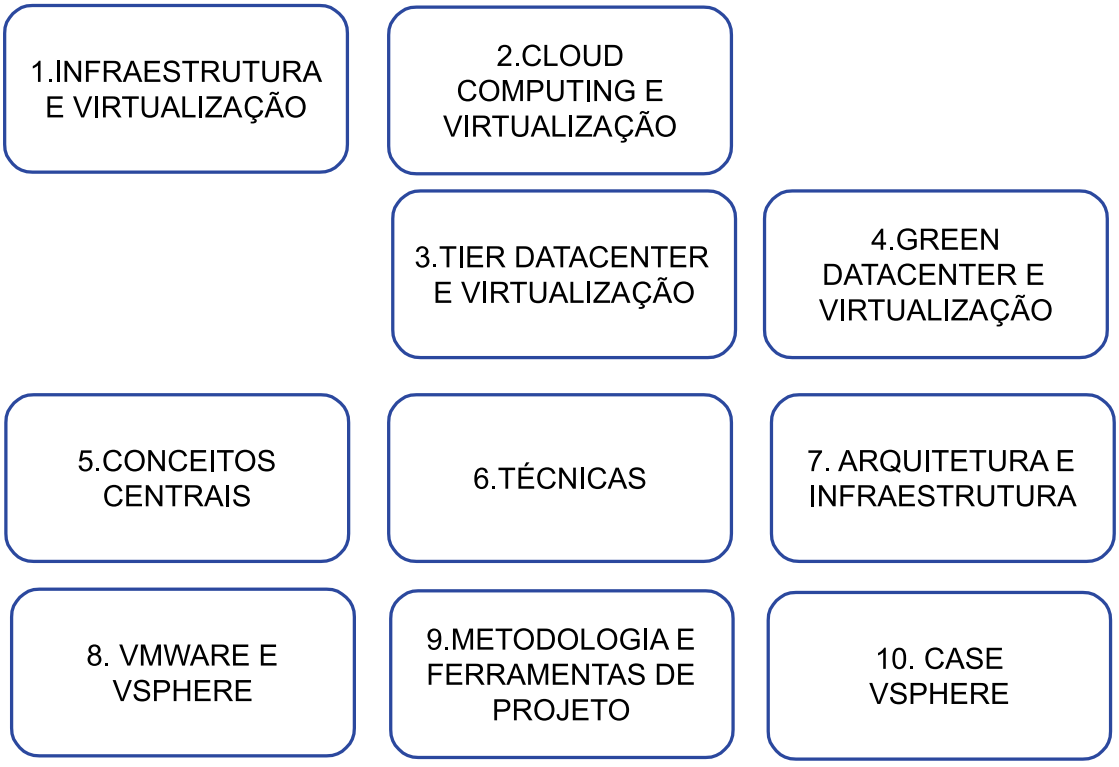


Figura 0-1 Capítulos do Livro

PARTE I: VISÃO GERAL

1. Infraestrutura de TI e Virtualização

1.1. Introdução

A infraestrutura de TI é o alicerce do modelo operacional da organização baseada na informação. O modelo operacional, conceito sugerido por Weill e Ross (2010), representa o nível de integração e padronização dos processos necessários ao bom funcionamento da organização. Por sua vez, o modelo operacional reflete a estratégia da organização.

A execução da estratégia, ancorada no modelo operacional, acaba dependendo da condição que a infraestrutura de TI proporciona. Alguns autores reforçam que a Infraestrutura de TI, no final das contas, é quem também responde pela condição de inovar de uma organização nos dias atuais, mesmo que no nível operacional[1].

Conceitualmente a infraestrutura de TI é a parte da TI que suporta as aplicações que, por sua vez, sustentam os processos de negócio. A **Figura 1-1** ilustra a relação entre infraestrutura de TI, aplicações e processos de negócio. Para alguns autores, como Weill e Ross, os recursos humanos da área de TI são considerados parte da infraestrutura de TI.

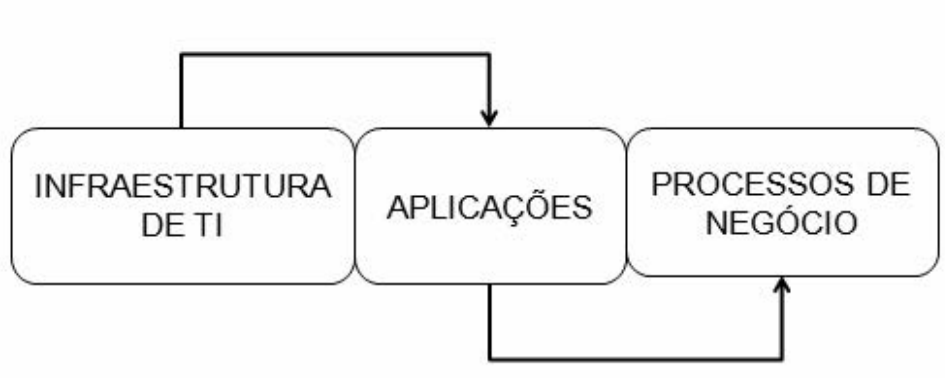


Figura 1-1 Infraestrutura de TI, Aplicações e Processos de Negócio

1 Ler o capítulo A Vantagem Essencial do livro O Futuro da Administração, de Gary Hamel e Bill Breen, publicado em 2008 no Brasil pela Elsevier.

A infraestrutura de TI também pode ser vista como o conjunto de serviços compartilhados de TI, disponível para toda a organização. O usuário da infraestrutura, na verdade, enxerga um serviço fornecido pela TI. A ideia central hoje é que o nível do serviço para o usuário, fornecido pela TI, é o que realmente importa.

A infraestrutura de TI, como qualquer outra infraestrutura, tem o papel de possibilitar que a organização funcione e cresça sem grandes interrupções. As organizações dependem cada vez mais da infraestrutura de TI, na medida em que trocam processos de negócios analógicos por processos digitais, que são a base do seu modelo operacional.

Outro conceito fundamental é o de governança de TI. A governança de TI deve alocar a responsabilidade pela definição, provisionamento e precificação destes serviços compartilhados de TI, buscando alinhar o nível destes serviços com as recomendações definidas na estratégia de TI para as aplicações. A estratégia de TI, por sua vez, deve estar de acordo com a estratégia da organização.

A governança de TI deve permitir a digitalização dos processos de forma consistente. O IDC[2] estima que, em 2020, o universo digital (toda informação criada e replicada em formato digital) será 44 vezes maior que em 2009, saindo de 0,8 ZB (1 ZB=1.000.000.000.000 GB) para 35 ZB. Pode-se assim ter uma ideia de como as organizações vão depender cada vez mais da infraestrutura de TI para operar. Cerca de 25% destes 0,8 ZB já são informação empresarial. A **Figura 1-2** ilustra o provável crescimento da base digital de informações, segundo o IDC.

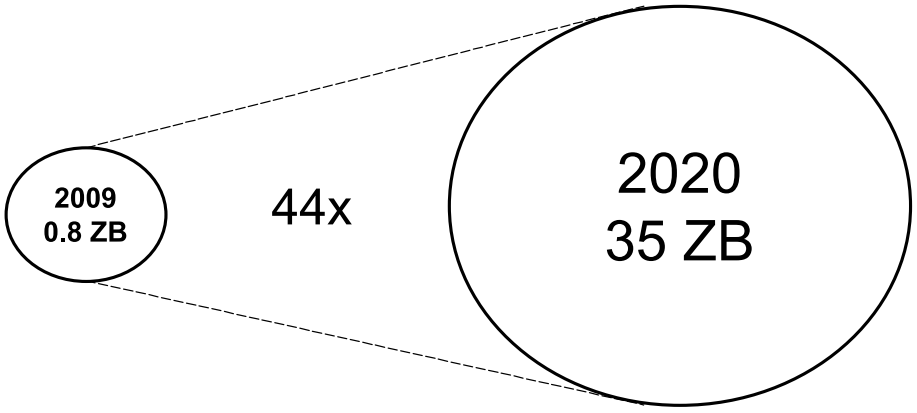


Figura 1-2 Expansão do Universo Digital (Fonte: IDC)

A introdução da banda larga em grande escala em vários países, incluindo o Brasil, reforça também a importância da infraestrutura de TI como alicerce importante do mundo baseado em informação. O acesso em banda larga é caracterizado pela disponibilização de infraestrutura de TI que possibilita tráfego de informações contínuo, ininterrupto e com capacidade suficiente para as aplicações de dados, voz e vídeo. Os Estados Unidos, por exemplo, definiram como marco o valor de 100Mbps como velocidade de conexão de *download* para cem milhões de residências americanas até 2020. O Brasil também já pos-

2 IDC – IVIEW The Digital Universe Decade – Are You Ready?, maio de 2010, por John Gantz e David Reinsel.

sui seu plano nacional de banda larga. O avanço da adoção da banda larga sinaliza a opção digital do mundo contemporâneo e reforça a necessidade de que governos e organizações privadas planejem a utilização de uma plataforma digital como fator de competitividade nacional.

Vale ressaltar que a infraestrutura de TI de hoje é mais complexa do que a infraestrutura de TI de alguns anos atrás, pois é uma combinação de infraestrutura privada (redes e dispositivos que conectam unidade de negócio, organização, setor de atuação) e pública (normalmente a Internet). A Internet é uma via pública e a garantia de serviços nesta rede é uma tarefa complexa. As opções referentes à infraestrutura de TI são muitas e as decisões precisam ser criteriosas, pois envolvem altos investimentos.

A dependência da organização da infraestrutura e das aplicações de negócio exige cada vez mais a participação dos gestores de TI em questões de planejamento e decisões de investimento. Esta participação normalmente encontra uma barreira em boa parte das organizações, pois normalmente é mais fácil para um executivo de alto escalão entender um investimento em marketing do que entender o investimento em TI. De qualquer forma, aos poucos, o gestor de TI vem aumentando o seu espaço dentro das organizações.

Dois componentes destacam-se como componentes fundamentais da infraestrutura de TI: o DATACENTER e a VIRTUALIZAÇÃO. O DATACENTER é o elemento central da infraestrutura de TI. É lá que os dados da organização são consolidados, processados e armazenados. Por sua vez, a VIRTUALIZAÇÃO permite otimizar o uso da infraestrutura de TI na medida em que torna mais inteligente a utilização dos recursos do DATACENTER.

1.2. Investimentos em Infraestrutura de TI

É preciso transformar a TI. Segundo Weill e Ross (2010), abordagens utilizadas para mudar a TI nos últimos anos se mostraram inadequadas. Entre elas destacaram-se:

- Pôr mais dinheiro nos problemas de TI: Em muitos casos, esta opção só aumentou os gastos e não os benefícios.
- Cortar drasticamente os gastos com TI: no curto prazo é uma saída que força o diálogo sobre as prioridades da empresa, mas pode minar a competitividade no longo prazo.
- Demitir o CIO: se for só para achar culpado, não resolve. O CIO, muitas vezes, não teve suas responsabilidades aceitas pelo restante da equipe administrativa e, portanto, não conseguiu exercer o seu papel.
- Terceirizar o problema da TI: pode não ser a solução se não houver mudança dos hábitos em relação a TI. Os custos e serviços possivelmente não melhorarão significativamente se as pessoas de negócio não modificarem os hábitos em relação a TI.
- Remover sistemas legados e substituí-los por um grande sistema integrado desenvolvido externamente (ERP ou coisa parecida): o sistema integrado resolve parte do problema, mas se não houver mudança na forma da gestão o sistema por si só não mudará a sorte da empresa.

Um dos conceitos chaves da transformação da TI é o conceito de alinhamento estratégico, que é o componente central da governança de TI e permite, quando bem feito, executar os projetos que são priorizados de acordo com a estratégia.

O alinhamento estratégico foca em garantir a ligação entre os planos de negócio e de TI, definindo, mantendo e validando a proposta de valor de TI, alinhando as operações de TI com as operações da organização. É comum existirem dois planos que se complementam: o plano estratégico de TI (PETI), que define os objetivos e projetos estratégicos, e o plano diretor de TI (PDTI), que trata do plano de execução dos projetos prioritários e da alocação de recursos em um horizonte de tempo.

Mas como fazer o tal alinhamento? Uma forma é utilizar o conceito de gestão de portfólio de projetos. Os projetos definidos pelas ações oriundas do Planejamento Estratégico precisam ser priorizados e escolhidos de forma transparente. A ideia é pensar os projetos como um portfólio e definir critérios de escolha que estejam alinhados com as prioridades da organização. A lógica atual do planejamento estratégico é partir do *Balanced Scorecard* (BSC), que traduz a estratégia, definir as ações e chegar aos projetos utilizando o conceito de portfólio de projetos. Utiliza-se a mesma lógica para a organização de TI.

Assim, no final, tem-se dois conjuntos de projetos: o portfólio de projetos da organização e o portfólio de projetos de TI, otimizados conforme ilustra a **Figura 1-3**. Neste ponto precisa-se alinhar os projetos de TI aos projetos da organização. Ou seja, os projetos de TI finalmente priorizados serão os que efetivamente suportam os projetos da organização. O Plano Diretor de TI (PDTI) seria o instrumento que normatiza a execução dos projetos priorizados de TI, tudo isto dentro de uma medida de tempo que pode ser, por exemplo, o ano fiscal.

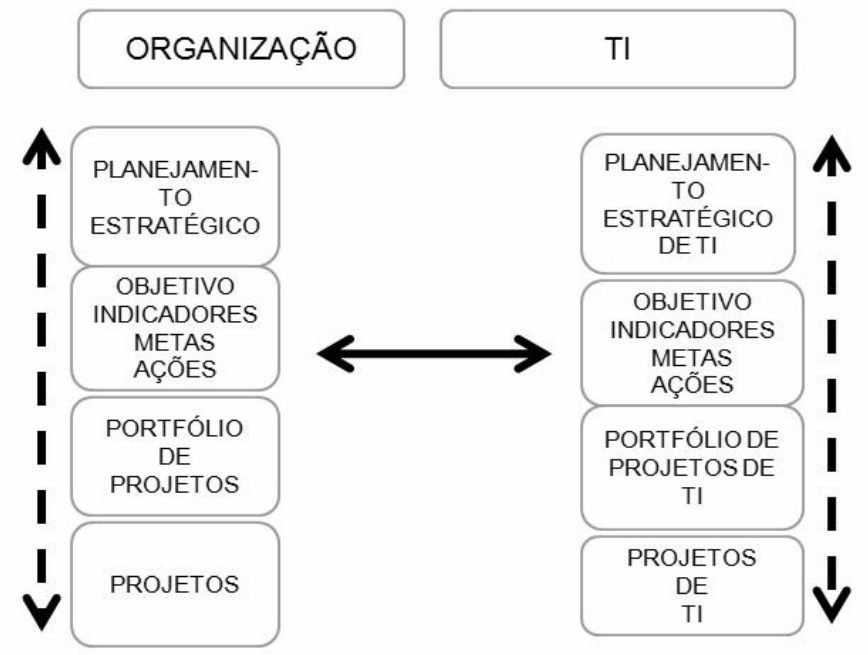


Figura 1-3 Alinhamento Estratégico

O conceito de portfólio de projetos de TI envolve a soma total dos investimentos em TI, incluindo aquisições de hardware, software, redes e contratação de pessoal. O portfólio de TI pode ser gerenciado como um porfólio financeiro, pesando-se riscos e retorno para o atingimento das metas empresariais. O alicerce do porfólio de TI é o investimento em infraestrutura de longo prazo.

Weill e Ross (2010) citam três componentes importantes para um novo modelo de financiamento da TI:

- Altos executivos estabelecem prioridades e critérios claros para os investimentos em TI.
- A gerência desenvolve um processo transparente para avaliar os projetos potenciais.
- Alocar os recursos e monitorar o impacto destas decisões de investimentos e usar o aprendizado para direcionar investimentos futuros.

Conforme mencionado anteriormente, um dos aspectos chaves da boa governança de TI é manter o tal alinhamento entre a estratégia do negócio e a estratégia da TI. Este alinhamento passa pela forma de como é gasto o dinheiro que vai para TI envolvendo investimentos e custeio. As organizações, de uma forma geral, estão tentando reduzir o custo da TI para que sobre mais dinheiro do orçamento para investimento em novos projetos. A organização de TI, hoje, faz um esforço de gastar cada vez menos com a sustentação da operação (o custeio) e mais com novas iniciativas (o investimento).

Sim, mas o que isto tem a ver com infraestrutura de TI? Decisões de investimento em infraestrutura de TI são muito importantes dentro das decisões a serem tomadas em TI. O que é novo em TI é que, agora, boa parte do investimento vai para infraestrutura. As organizações dependem da infraestrutura para operarem internamente, se integrarem a bancos, seguradoras, governo, clientes e fornecedores.

Segundo Weill e Broadbent (2004), o investimento em infraestrutura de TI já representava, em 2004, cerca de 60% dos investimentos das empresas em TI nos Estados Unidos. A consultoria McKinsey&Company também realizou uma pesquisa em grandes empresas americanas que sugeriu que 25% do orçamento de TI de grandes empresas dos Estados Unidos vai para o DATACENTER (DC). A pesquisa aponta que, dos 60% dos investimentos em infraestrutura, 25% vai para o DATACENTER, o que o torna o elemento mais importante da infraestrutura. O DATACENTER é onde boa parte dos dados e informação são processados e armazenados. A **Figura 1-4** ilustra os resultados da pesquisa realizada pela consultoria.

Se uma grande empresa no Brasil tem orçamento de R$ 2M para a TI, existe uma boa chance de gastar mais de R$ 1M com infraestrutura. Como deixar esta decisão na mão de alguém que não entende a estratégia da organização? Como definir os investimentos em infraestrutura? Priorizar o quê? A rede de longo alcance WAN? As redes locais LAN? Os servidores ou a segurança? É necessário ter uma clara visão de como definir estes investimentos e como transformar estas aquisições em bons indicadores baseados em serviços quando da operação da nova infraestrutura. Os novos investimentos devem vir para melhorar os níveis de serviço fornecidos e possibilitar o crescimento da organização.

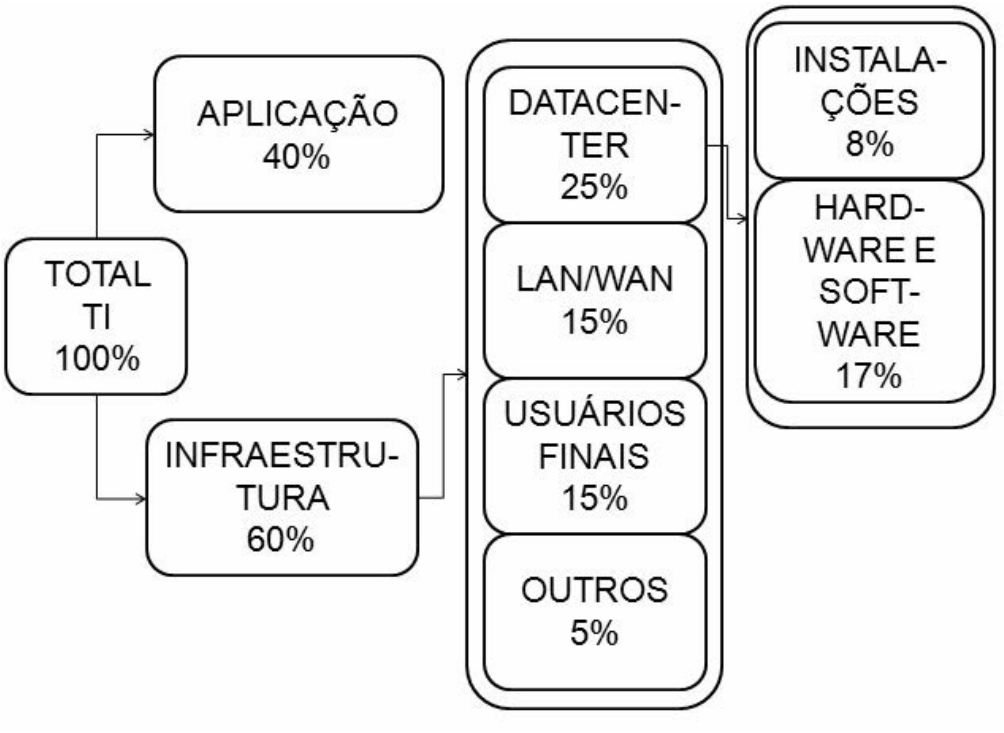


Figura 1-4 Distribuição do Orçamento de TI (Fonte: McKinsey&Company)

Os aspectos citados acabam sendo os sugeridos pelo *framework* COBIT e aqui resumidos:

- Entrega de Valor: garantir que a TI entregue os benefícios previstos na estratégia da organização.
- Gestão de Recursos: refere-se à melhor utilização dos investimentos e gerenciamento dos recursos de TI incluindo informações, aplicativos, infraestrutura e pessoas.
- Gestão de Riscos: trata de dar transparência aos riscos significantes para a organização e do seu gerenciamento, além de cuidar dos requisitos de conformidade.
- Mensuração de Desempenho: trata de acompanhar e monitorar a implementação da estratégia de TI.

1.3. Maturidade da Infraestrutura de TI

1.3.1. Introdução

Uma maneira consagrada de verificar a situação de um determinado processo ou uso de uma determinada tecnologia é verificar a situação de uso em termos de sua maturidade. Normalmente o nível de maturidade encontrado é comparado com um padrão e definem-se ações que permitem melhorar a situação encontrada. Esta é uma possível estratégia de alterar o status de uso de uma determinada tecnologia ou processo. Pode-se pensar desta forma também para a infraestrutura de TI. Ou seja, uma maneira de

pensar a infraestrutura de TI é compará-la com um padrão e depois definir ações que a tornem pelo menos aderente à realidade da organização, o que se convencionou chamar de alinhamento.

Conforme dito anteriormente, a execução do modelo operacional que reflete a estratégia da organização depende do suporte da infraestrutura de TI. É necessário otimizar a utilização da infraestrutura de TI, pois a empresa depende cada vez mais dela.

1.3.2. Modelo de Weill e Broadbent

Weill e Broadbent (1998) constataram, em pesquisa realizada nos Estados Unidos, que as organizações normalmente aderem a uma das quatro visões de infraestrutura de TI mencionadas abaixo:

- **Nenhuma**: investimento baixo quando comparado à concorrência, 0% de investimento em infraestrutura comum na empresa, abordagem para investimento sem justificativa, nenhuma extensão dos serviços de infraestrutura.
- **Utilitária**: investimento baixo quando comparado à concorrência, investimento em infraestrutura comum abaixo da média – 37% da TI total, justificativa de investimento por custo, extensão dos serviços dentro e entre unidades de negócio para dados e transações simples, extensão de serviços básica.
- **Dependente**: investimento médio quando comparado com à concorrência, investimento em infraestrutura comum levemente acima da média – 45% da TI total, justificativa de investimento baseado em equilíbrio entre custo e flexibilidade, extensão dos serviços dentre e entre unidades de negócio, algumas transações complexas e extensão básica com alguns serviços complexos.
- **Facilitadora**: investimento mais alto de que o da concorrência, investimento em infraestrutura comum muito acima da média – 50% ou mais da TI total, abordagem de justificativa de investimento pela flexibilidade, extensão dos serviços dentro e entre unidades de negócio, algumas transações complexas, extensão abrangente.

Uma forma de transformar a infraestrutura de TI é diagnosticar a situação da infraestrutura de TI utilizando o modelo citado e definir ações que permitam transformá-la melhorando o seu nível de maturidade. Estas ações deveriam estar contempladas no Planejamento Estratégico da TI (PETI).

Os autores citados também salientam que a capacidade da infraestrutura da TI pode ser definida segundo dois fatores:

- Serviços de infraestrutura oferecidos ao longo da empresa;
- Alcance e escopos eletrônicos que se referem às localidades e às pessoas que podem ser interligadas.

Estes dois fatores podem ser utilizados como indicadores de maturidade da infraestrutura de TI.

1.3.3. Modelos da Microsoft

A Microsoft desenvolveu o Modelo de Otimização de Infraestrutura (*Infrastructure Optimization Model* – IOM) usando práticas recomendadas pela indústria, por instituições como o CISR do MIT e a sua própria experiência com os clientes. O modelo proposto pela Microsoft oferece uma estrutura para identificar as pessoas, os processos e a tecnologia envolvidas na infraestrutura de TI e descreve quatro níveis de maturidade através dos quais esses recursos de TI podem evoluir para aumentar cada vez mais o valor estratégico da infraestrutura para a organização.

O modelo IOM oferece um ponto de partida para que as organizações avaliem o estado atual de sua infraestrutura de TI e aprendam a alcançar o nível de maturidade apropriado para seus negócios. Usando o modelo, os clientes entendem rapidamente o valor estratégico e os benefícios de mudar de um uso "básico" de infraestrutura, onde a infraestrutura da TI geralmente é considerada um centro de custos, para um uso "dinâmico", onde a infraestrutura de TI é vista como um bem estratégico que contribui significativamente para o desempenho dos negócios.

Os diferentes níveis do modelo e a forma que a TI é vista na organização podem ser caracterizados de acordo com a **Figura 1-5**.

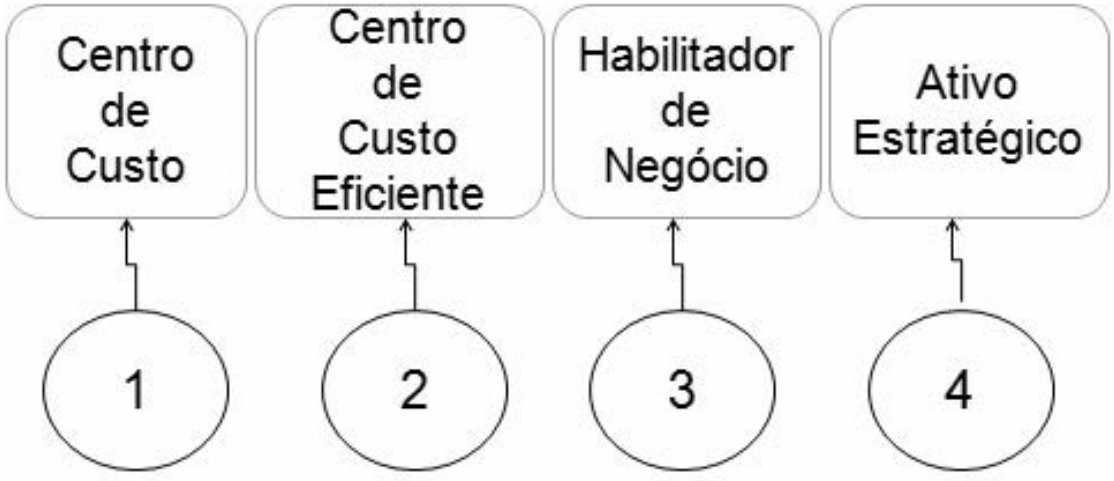


Figura 1-5 Modelo IOM

1. **Básico:** processos manuais; nenhum controle central; segurança reativa; administração e serviços com alta intervenção e alto custo para usuários. A infraestrutura básica é caracterizada por processos manuais, localizados, controle central mínimo, políticas e padrões de TI não existentes ou não executados em relação a segurança, backup, gerenciamento e implantação de imagens, conformidade e outros padrões de TI comuns. Há uma falta geral de conhecimento em relação aos detalhes da infraestrutura que está sendo utilizada no momento ou às práticas que terão maior impacto em sua melhoria. O funcionamento geral das aplicações e serviços é desconhecido devido à falta de ferramentas e recursos. Não há veículo para compartilhar o conhecimento acumulado ao longo do ciclo de vida da

TI. Organizações com infraestrutura básica acham seus ambientes extremamente difíceis de controlar, têm custos muito altos com gerenciamento de servidor e desktops, geralmente são muito reativos a ameaças de segurança e têm pouco impacto positivo sobre a capacidade de o negócio se beneficiar do TI. Normalmente todas as atualizações, implantações de software e serviços são fornecidas por meio de alto custo e alta interação.

2. **Padronizado**: processos possuem alguns padrões e diretivas; o controle é central; uso limitado de automação; ainda é reativa na segurança; administração e serviços com intervenção e custo médios para usuários. A infraestrutura padronizada introduz controles por meio da utilização de padrões e políticas para gerenciar desktops e servidores, sobre como máquinas são introduzidas na rede, da utilização do *Active Directory* (AD) para gerenciar recursos, políticas de segurança e controles de acesso. Clientes em um estado padronizado compreendem o valor de padrões básicos, e algumas políticas ainda são bem reativas. Normalmente todas as atualizações, implantações de software e serviços de desktops são fornecidos através de média interação com médio ou alto custo. Entretanto, estas organizações têm um inventário razoável de hardware e software e estão começando a gerenciar licenças. Medidas de segurança são aprimoradas com um perímetro fechado, mas a segurança interna ainda pode ser um risco.

3. **Racionalizado**: processos controlados centralmente; segurança proativa; administração e serviços sem intervenção para usuários; diretivas de TI estão começando a dar suporte e expandir os negócios. A infraestrutura racionalizada é onde os custos envolvidos no gerenciamento de desktops e servidores estão em um nível mínimo, e processos e diretivas foram otimizados, a fim de começarem a desempenhar um papel importante no suporte e na expansão dos negócios. A segurança é bastante proativa, e a reação a ameaças e desafios é rápida e controlada. O uso da implantação totalmente automatizada ajuda a minimizar o custo, o tempo de implantação e os desafios técnicos. O número de imagens é mínimo, e o processo de gerenciamento de *desktops* requer muito pouco trabalho manual. Estas organizações têm um inventário claro do hardware e software e compram somente as licenças e os computadores de que precisam. A segurança é extremamente proativa, com diretivas e controle rígidos, desde o *desktop* até o servidor, o *firewall* e a *extranet*.

4. **Dinâmico**: processos automatizados, administração e serviços totalmente automatizados para os usuários; contratos de nível de serviço vinculados aos negócios; habilidade de adotar e inovar rapidamente com novas tecnologias para satisfazer necessidades estratégicas. As organizações com uma infraestrutura dinâmica estão totalmente cientes do valor estratégico que suas infraestruturas de TI fornecem para ajudá-los a fazer negócios com mais eficiência e a ficar na dianteira dos concorrentes. Os custos são controlados; existe uma integração entre usuários e dados, desktops e servidores; a colaboração entre usuários e departamentos é generalizada; e os usuários móveis têm quase os mesmos níveis de serviços e recursos que teriam no local, independentemente de onde estejam. Os processos de negócio e a TI são uma entidade única, o que permite que a equipe de TI esteja

alinhada com as necessidades comerciais. Investimentos adicionais em tecnologia geram benefícios mensuráveis específicos e rápidos para a empresa. O uso de softwares que se autoconfiguram e de sistemas semelhantes à quarentena para garantir o gerenciamento e a conformidade de patches com diretivas de segurança estabelecidas permite à organização com uma infraestrutura dinâmica automatizar processos e, assim, ajudar a aumentar a confiabilidade, reduzir os custos e melhorar os níveis de serviço.

Outro modelo da Microsoft trata da otimização da infraestrutura básica, denominado modelo CORE IO. O modelo CORE IO ajuda a organização a entender e avançar em direção a uma infraestrutura de TI mais segura, mais bem gerenciada e dinâmica, ajudando a reduzir o custo total de propriedade da TI, fazendo melhor uso dos recursos e transformando a TI em um ativo estratégico. O modelo CORE IO suporta profissionais de TI no gerenciamento de servidores, *desktops*, dispositivos móveis e aplicações, bem como no uso eficiente de recursos para ajudar a eliminar complexidades e custos desnecessários, garantir que a organização esteja sempre funcionando e estabelecer uma infraestrutura com maior capacidade de reação. O modelo de otimização CORE IO inclui capacidades técnicas específicas que fornecem um conjunto detalhado de soluções para ajudar a avançar a infraestrutura e a plataforma de otimização de um nível para o seguinte.

O modelo CORE IO define então cinco capacidades que são necessárias para construir uma infraestrutura de TI mais ágil:

- **Gerenciamento de Identidades e Acesso.** Esta capacidade descreve maneiras para gerenciar pessoas e identidades de ativos, bem como soluções que devem ser implementadas para gerenciar e proteger dados de identidade (sincronização, gerenciamento de senhas e provisionamento de usuário, apenas para mencionar alguns), além de descrever como gerenciar acesso a recursos para usuários corporativos móveis e parceiros do lado de fora de um *firewall*.
- **Gerenciamento de Estação de Trabalho, Dispositivo e Servidor.** Esta capacidade descreve como gerenciar *desktops*, dispositivos móveis e servidores, assim como implantar atualizações, sistemas operacionais e aplicativos na rede. Ela também fornece orientação sobre como se pode utilizar tecnologias de VIRTUALIZAÇÃO e de escritório remoto para melhorar a infraestrutura de TI.
- **Segurança e Redes.** Esta capacidade descreve o que deve ser levado em conta ao implementar a infraestrutura de TI para ajudar a garantir que as informações e as comunicações estejam protegidas contra o acesso não autorizado. Esta capacidade também descreve mecanismos para proteger a infraestrutura de TI contra ataques de negação de serviço e vírus, ao mesmo tempo em que preserva o acesso a recursos corporativos.
- **Proteção e Recuperação de Dados.** Esta capacidade fornece orientações para gerenciamento de restauração, armazenamento e *backup* estruturados. À medida que o armazenamento de dados e de informações se prolifera, as organizações ficam cada vez mais sob pressão para oferecer proteção e recuperação eficiente e com boa relação custo-benefício. Esta capacidade fornece orientações nesta área.

- **Processo de TI e Segurança.** Esta capacidade fornece orientações para práticas recomendadas sobre como projetar, desenvolver, operar e dar suporte a soluções de maneira eficaz em termos de custo, ao mesmo tempo em que alcança alta confiabilidade, disponibilidade e segurança. Apesar de uma tecnologia sólida ser necessária para atender às demandas por serviços de TI altamente seguros, confiáveis e disponíveis, a tecnologia em si não é suficiente. É preciso oferecer excelência em processos e pessoas (funções, habilidades e responsabilidades).

A ideia do modelo CORE IO é melhorar estas cinco capacidades tendo como pano de fundo o modelo IOM, conforme ilustra a **Figura 1-6**.

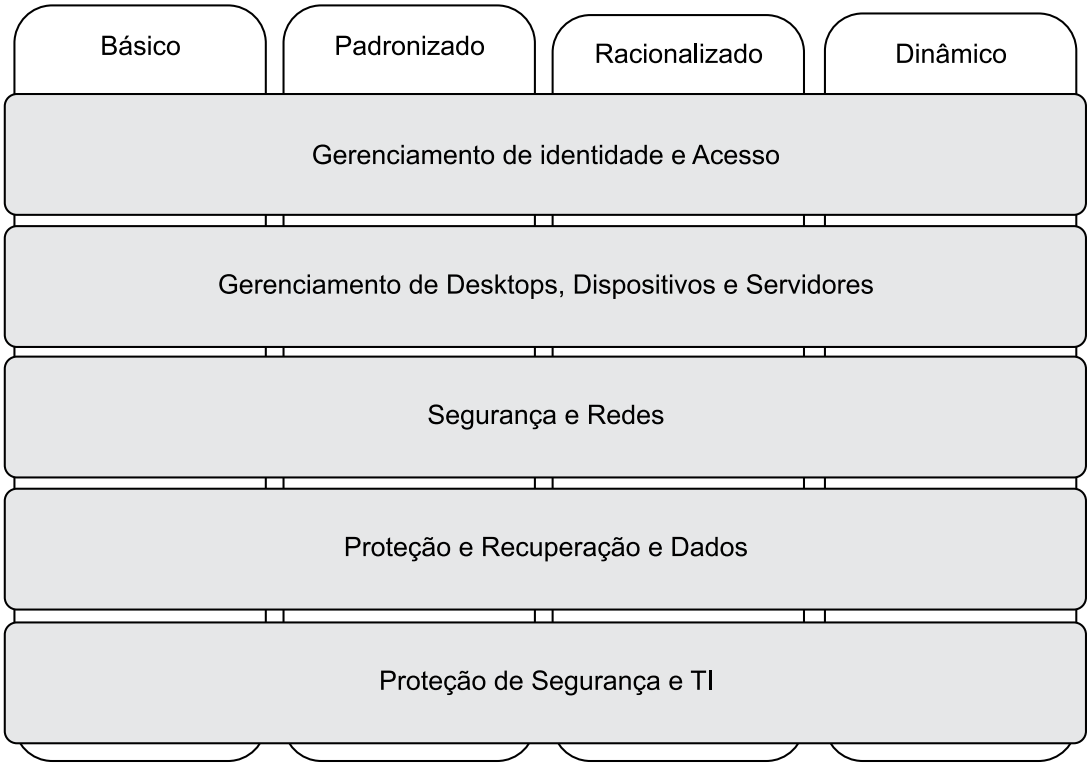


Figura 1-6 Modelo CORE IO

1.4. Gerenciamento de Serviços de TI

Na década de 80, a qualidade da infraestrutura de TI e, consequentemente, dos serviços oferecidos pelas unidades de TI do governo britânico era precária. O próprio governo solicitou a então CCTA (*Central Computer and Telecommunications Agency*) que desenvolvesse uma nova forma de trabalho dentro das organizações de TI que permitisse obter melhores resultados com um melhor uso dos recursos e com custos menores.

A ideia central do governo britânico era otimizar o uso dos recursos alterando o foco e saindo da opção de comprar equipamentos de TI para a opção de adquirir serviços de TI. De nada adiantaria ter uma infraestrutura complexa e cara se os serviços de TI

ofertados para os usuários eram precários. Vincular a infraestrutura de TI aos serviços de suporte e entrega fornecidos pela equipe de TI era a questão central. Ao mesmo tempo, deveria-se cuidar dos processos que permitiam entregar estes serviços de uma forma consistente.

1.4.1. ITIL

Este esforço resultou na criação da Biblioteca de Infraestrutura de Tecnologia da Informação (*Information Technology Infrastructure Library – ITIL*), desenvolvida entre 1989 e 1995, que consistia inicialmente de 31 livros que cobriam todos os aspectos do gerenciamento de serviços de TI. A ITIL atualmente esta sob custódia da OGC (*Office for Government Commerce) e* foi desenvolvida a partir de uma coleção de melhores práticas observadas no setor de infraestrutura de TI. Diversas outras bibliotecas foram desenvolvidas por organizações comerciais baseadas na ITIL, o que a tornou um padrão de fato para o gerenciamento de serviços de TI. A ITIL permitiu pensar a infraestrutura de TI em termos de serviços gerenciados. A qualidade dos serviços depende dos processos internos e das funções da TI, conforme ratifica a ITIL.

A segunda versão da ITIL (ITIL v2) foi estruturada entre 2000 e 2004 e constou de sete livros. A ITIL v2 se tornou um padrão e foi incorporada a norma BS 15000, que se tornou um anexo da norma ISO 20000.

Na prática, é comum a organização que pretende adotar a ITIL como padrão para a gerência de serviços priorizar inicialmente a implantação de algumas práticas de serviços em detrimento de outras, devido à complexidade e à mudança inicial envolvida. Ou seja, normalmente a implantação de uma biblioteca como a ITIL se dá de forma gradativa e a opção inicial é a de priorizar as práticas que estão mais urgentes, considerando a necessidade do negócio.

A ITIL também permite pensar de maneira integrada toda a cadeia de prestação de serviços (*end-to-end services*) e deixar de lado qualquer visão fragmentada de serviços.

1.4.2. ITIL v3

Em 2007, surgiu a versão três da ITIL (ITIL v3) que, no seu núcleo, consiste de cinco livros cobrindo o ciclo de vida dos serviços. Este é o principal avanço desta versão: tratar serviços sob o ponto de vista de ciclo. Existe também um sexto livro que faz uma introdução oficial ao ITIL e descreve o material abordado nos outros cinco livros. A **Figura 1-7** ilustra a estrutura de publicações da ITIL v3.

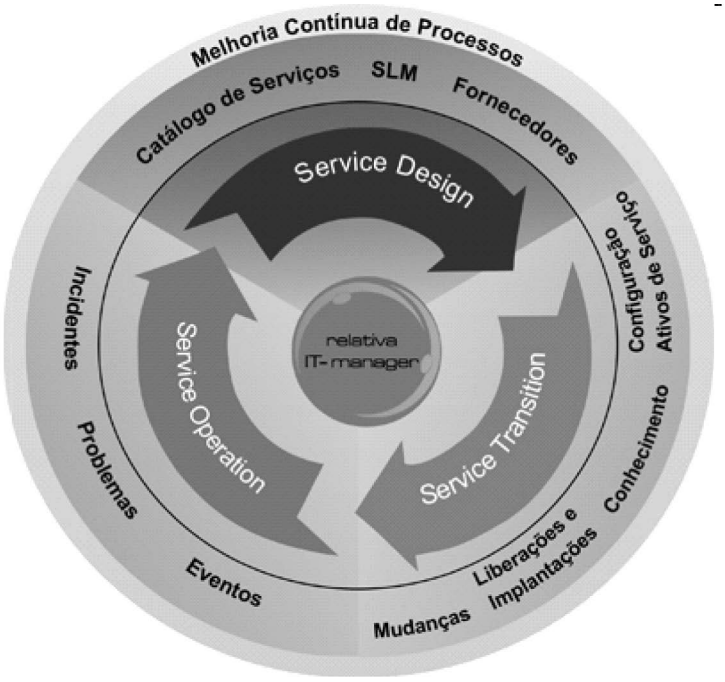


Figura 1-7 ITIL v3

A ITIL v3 três também traz um conjunto complementar de publicações com orientações específicas para setores da indústria, tipos de organização, modelos operacionais e arquiteturas de tecnologia.

Os cinco livros da versão 3 cobrem cada estágio do ciclo de vida dos serviços descritos a seguir:

- **Estratégia de serviço.** Transformar os serviços de TI em ativos estratégicos para atender os objetivos estratégicos da organização.
  - Processos e Funções
    - Processo de Gerenciamento Financeiro.
    - Processo de Gerenciamento de Portfólio de Serviço.
    - Processo de Gerenciamento da Demanda.
    - Função de Gerenciamento do Relacionamento com o Negócio.
- **Desenho de serviço.** Orientar a concepção dos serviços de TI para garantir a qualidade, a satisfação do cliente e a relação custo-benefício na prestação dos serviços.
  - Processos
    - Processo de Gerenciamento do Nível de Serviços.
    - Processo de Gerenciamento do Catálogo de Serviços.
    - Processo de Gerenciamento da Capacidade.
    - Processo de Gerenciamento da Continuidade dos Serviços de TI.

- Processo de Gerenciamento da Segurança da Informação.
- Processo de Gerenciamento de Fornecedores.

- **Transição de serviço**. Orientar o desenvolvimento de recursos para implementação de serviços novos ou modificados na operação de TI e garantir que os serviços definidos pela estratégia de serviços e planejados no desenho dos serviços estão sendo efetivamente realizados na operação para controlar e minimizar riscos de fracasso ou ruptura dos serviços.
  - Processos
    - Processo de Planejamento e Suporte de Transição.
    - Processo de Gerenciamento de Mudanças.
    - Processo de Gerenciamento da Configuração e de Ativos de Serviço.
    - Processo de Avaliação.
    - Processo de Validação e Teste de Serviço.
    - Processo de Gerenciamento do Conhecimento.
- **Operação do serviço**. Orientar sobre como alcançar a eficiência e a eficácia na entrega e suporte dos serviços para garantir o valor esperado pelo cliente e o atendimento dos objetivos estratégicos da empresa.
  - Processos e Funções
    - Processo de Gerenciamento de Eventos.
    - Processo de Gerenciamento de Incidentes.
    - Processo de Cumprimento de Requisição.
    - Processo de Gerenciamento de Problemas.
    - Processo de Gerenciamento de Acesso.
    - Função Central de Serviços.
    - Função Gerenciamento Técnico.
    - Função Gerenciamento da Operação de TI.
    - Função Gerenciamento de Aplicações.
- **Melhoria continuada de serviços**. Identificar resultados e orientar sobre a melhoria de serviços unindo esforços com os ciclos de Estratégia, Desenho, Transição e Operação para criar ou manter o valor dos serviços.
  - Escopo
    - Melhoria dos Ciclos de Vida dos Serviços.
    - Melhoria do alinhamento do Portfólio de TI com os requerimentos atuais e futuros e os Serviços de TI em produção.
    - Aumento da maturidade das entregas dos processos em cada serviço de TI.

Pesquisa realizada pela FGV (GVcia), em 2007, verificou a intensidade do uso de práticas de gerenciamento de serviços de TI no Brasil. O estudo sugere o amplo domínio da ITIL (50,4%) quando se trata do uso de metodologias de gerenciamento de serviços de TI reconhecidas pelo mercado.

1.4.3. TCO

O conceito de custo total de propriedade (*Total Cost of Ownership – TCO*) foi adaptado pelo Gartner, em 1987, para aplicação na área de TI como um meio de entender os custos reais da infraestrutura de TI. Na época existia uma grande demanda em medir a melhoria ou mesmo a perda de produtividade oriunda do uso indiscriminado de redes de computadores baseadas em servidores e *desktops*. O conceito de TCO já era conhecido e bastante utilizado em outras áreas. O Gartner então transferiu e adaptou o conceito para a TI. Todos achavam que o custo de manter estas redes era muito maior do que se imaginava, mas quase ninguém sabia como medir estes custos, pois existiam aspectos subjetivos de difícil mensuração.

A ideia do uso do TCO na área de TI foi a de medir o custo total ao longo do ciclo de vida da solução tecnológica e não só o custo de aquisição. Os fornecedores de hardware de boa qualidade rapidamente se adaptaram ao novo conceito sempre tentando provar para os clientes que agora não importava mais só o custo de aquisição de um determinado produto e que era preciso avaliar os custos durante pelo menos o período de garantia, normalmente de três a cinco anos. O conceito de TCO está ligado diretamente ao conceito de ROI (*Return on Investiment*). ROI é a relação entre o dinheiro ganho ou perdido através de um investimento e o montante de dinheiro investido.

O TCO tem dois principais componentes:

- **Custos Diretos**: tradicionalmente são os mais fáceis de mensurar e, por esse mesmo motivo, frequentemente recebem uma atenção maior. Exemplos de custo direto são: hardware, software, operação e administração.
- **Custos Indiretos**: são mais vagos e difíceis de mensurar. Por sua natureza "invisível", os custos indiretos são muitas vezes subestimados. Exemplos de custo indireto são o custo com o downtime e o custo com o usuário final. Na verdade, o custo direto influencia o custo indireto. Se o hardware adquirido é de baixa qualidade, o custo direto influencia o custo indireto, pois o custo com o downtime, por exemplo, tende a ser maior.

Outro aspecto relevante é que os custos diretos são os custos possíveis de serem orçados e os custos indiretos são os custos não orçados, pois não são visíveis ao orçamento.

O modelo de TCO desenvolvido pelo Gartner para TI agrupa os custos em quatro categorias (planejamento, aquisição, operação e manutenção, alienação). O Gartner logo nas primeiras pesquisas constatou que os custos com o usuário final, ou seja, os custos de operação e manutenção eram maiores do que os custos de aquisição de hardware e software. Começava-se a mostrar com números a ineficiência do modelo cliente-servidor.

Atualmente, diversos estudos tentam comparar o TCO entre produtos na área de TI. Um bom exemplo da importância do conceito é a quantidade de estudos realizados que estimam e comparam o TCO para os Sistemas Operacionais Windows e para o Linux . Também existem estudos com foco no TCO para aplicações específicas como *Contact Centers*, Computação Distribuída e para DATACENTERS. Também há uma proliferação de estudos que normalmente tentam demonstrar que uma determinada tecnologia, quando utilizada corretamente, permite reduzir o TCO do ambiente de TI.

A redução do TCO é uma prioridade da maioria das organizações e um modelo para avaliação do TCO deve permitir identificar e mensurar custos e benefícios, visando reduzir os custos e aperfeiçoar os benefícios. A simplificação da Infraestrutura de TI é um dos principais pilares para a redução do TCO de uma organização. Quanto mais as organizações dependem da infraestrutura, maior deve ser o esforço para torná-la simples e gerenciável, atributos fundamentais na busca da redução do TCO.

1.4.4. ITIL e TCO

As duas principais abordagens utilizadas com foco em infraestrutura de TI, o modelo do TCO e a biblioteca ITIL meio que se completam. Os clientes demandam melhoria nos serviços de TI para atender as expectativas de negócio e as organizações utilizam a ITIL para melhorar a entrega e o suporte dos serviços de TI.

Por sua vez, a alta direção cobra que os custos de TI sejam justificados, gerenciados e até reduzidos. Neste caso, a abordagem de TCO permite trabalhar o custo durante todo o ciclo de vida do serviço e assim é utilizada para ajudar a reduzi-lo. Desta maneira, o TCO apoia a ITIL, quantificando os custos diretos e alocando-os corretamente no orçamento, considerando também os custos indiretos, possibilitando a utilização da ITIL para o atingimento da melhoria dos serviços de TI.

1.5. Impactos da VIRTUALIZAÇÃO na Infraestrutura de TI

O ritmo das mudanças no cenário globalizado exige das organizações uma maior flexibilidade para inovar e, portanto, a manutenção de uma infraestrutura de TI cada vez mais adaptável. As mudanças alteram os processos organizacionais que, por sua vez, alteram as necessidades de infraestrutura de TI. Por exemplo, determinada empresa acaba de ser adquirida por outra. A rapidez necessária para aperfeiçoar e adequar os novos processos das duas organizações, agora transformadas em uma, exige que a infraestrutura de TI esteja pronta para a mudança.

As organizações buscam construir plataformas digitais mais estáveis que permitam o crescimento mesmo em ambientes turbulentos. Parece contraditório mas não é, o que se quer é o melhor dos dois mundos: plataformas estáveis e flexíveis. Pode-se pensar em arquiteturas e infraestrutura que suportem o crescimento do negócio. A VIRTUALIZAÇÃO, neste contexto, é uma peça chave, pois permite alterar a infraestrutura rapidamente utilizando instrumentos lógicos e não físicos, ao mesmo tempo que estabiliza o ambiente tornando as aplicações independentes do hardware.

Assim, a VIRTUALIZAÇÃO desvincula as aplicações e o sistema operacional dos recursos físicos e acaba por agilizar e permitir o surgimento da plataforma digital. Servidores virtuais são mais fáceis de serem instalados, gerenciados e migrados para o novo ambiente. Também os ambientes de teste e desenvolvimento, necessários para a manutenção de aplicações empresariais, são facilmente montados com a otimização dos recursos acarretada pela VIRTUALIZAÇÃO.

Pesquisas do IDC revelam que apenas 15% da capacidade dos servidores é usada nas empresas. Os 85% restantes estão ociosos. Por essa razão, é importante utilizar a VIRTUALIZAÇÃO para otimizar o uso destes recursos. A consolidação de servidores e o impacto causado vai desde aperfeiçoar o uso do espaço físico e melhorar o gerenciamento do ambiente de TI até a redução dos custos com o consumo de energia.

O modelo dinâmico de infraestrutura de TI proposto pela Microsoft e citado anteriormente só é possível de ser construído se existe uma infraestrutura de TI adaptável. A VIRTUALIZAÇÃO é a tecnologia utilizada na busca de uma infraestrutura de TI que se autoconfigura de acordo com a demanda das aplicações e dos negócios. As aplicações, de uma forma geral, estão também se adaptando ao utilizar o modelo orientado a serviço e o desenvolvimento rápido. Agora é a vez da infraestrutura de TI ser construída de forma a se adaptar às demandas dos negócios. A VIRTUALIZAÇÃO é o elemento central do novo DATACENTER.

1.6. Referências Bibliográficas

An Introductory Overview of ITIL v3. itSMF, 2007.

Cobit 4.1, Tradução. IT Governance Institute, 2009.

Freitas, André dos Santos. Fundamentos do Gerenciamento de Serviços de TI. Brasport, 2010.

IT Value Transformation Roadmap: Vision, Value and Virtualization. IT Process Institute, 2010.

Leveraging Microsoft Optimization to Create Your Dynamic Roadmap. Microsoft, 2009.

Opções atuais para melhoria do custo total de propriedade (TCO). Tradução. EMC, maio de 2005.

Opções atuais para simplificar a infraestrutura de informações. Tradução. EMC, maio de 2005.

Revolutionizing DATACENTER Energy Efficiency. McKinsey&Company, jul. 2008.

Simplify and Transform Your IT Infrastrcture into a Strategic Business Asset. Microsoft & Dell, 2009.

Sturm, Rick. Service Level Management: Fundamentos do Gerenciamento de Níveis de Serviço. Tradução. Editora Campus, 2000.

Veras, Manoel. DATACENTER: Componente Central da infraestrutura de TI, Brasport, 2009.

Virtualização: Eficiência sob Medida. Computerworld, Executive Briefing.

Weill, Peter & Broadbent, Marianne. Leveraging the New Infrastructure: How Market Leaders capitalize on Information Technology. HBR School Press, 1998.

Weill, Peter & Ross, Jeanne W. Conhecimento em TI. Tradução. M.Books, 2010.

Weill, Peter & Ross, Jeanne W. Governança de TI. Tradução. M.Books, 2006.

2. Cloud Computing e Virtualização

2.1. Introdução

As empresas estão se organizando em formato de rede. Os processos de negócio entre estas organizações se utilizam cada vez mais das aplicações que processam e fornecem as informações necessárias para o pleno funcionamento deste novo arranjo. Por sua vez, a TI é a "cola" que permite que estas organizações trabalhem em conjunto com um determinado objetivo e entreguem para o cliente final um valor que somado é maior do que o que se conseguiria com as partes isoladas sem a TI. Ou seja, a TI é que viabiliza a organização em rede.

Esta nova organização, diferente da empresa vertical de Ford, é baseada em arranjos horizontais. A ideia da rede é ser uma opção para mercados e hierarquia e o formato em rede se beneficia da redução dos custos de transação propiciada pela TI.

A **Figura 2-1** ilustra a referida mudança. A TI, no caso da organização com forma de pirâmide, era de certa forma centralizada e exclusiva, mesmo que os serviços fossem prestados por terceiros. Com o formato em rede a TI assumiu a forma descentralizada.

A rede inicialmente era interna, ou seja, lá trafegavam dados e informações de uma mesma organização. Formar redes entre organizações era uma tarefa complexa e cara. Existiam quase sempre intermediários que comandavam a interligação entre as redes e cobravam caro por isto. Além disso, os padrões e protocolos utilizados eram muito específicos de cada setor.

O surgimento da INTERNET e a consequente redução dos custos de interligação, o avanço da padronização de protocolos de comunicação e a disponibilidade de banda que ocorreram muito rapidamente. Um efeito positivo do exagero com os negócios na INTERNET, por volta do ano 2000, possibilitaram tornar a opção de interligar redes de diferentes organizações uma realidade. Sistemas interorganizacionais são agora baseados nestas redes que utilizam a INTERNET como espinha dorsal. As fronteiras organizacionais não são claras neste novo modelo organizacional, e a TI acaba por deixar mais difícil saber quais os limites existentes entre essas organizações.

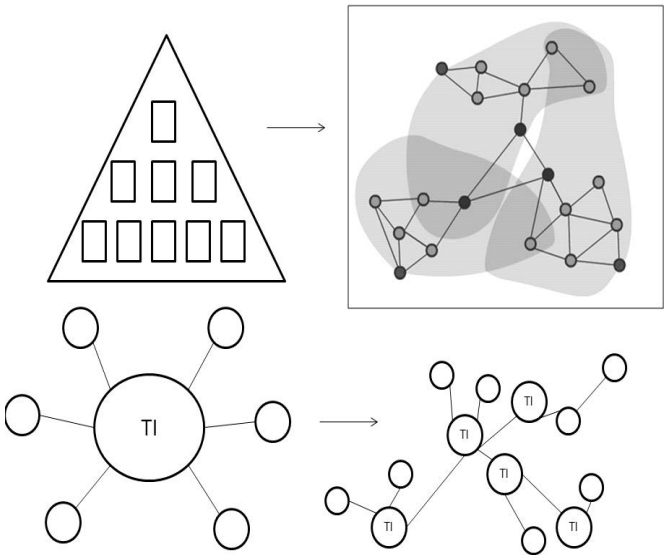
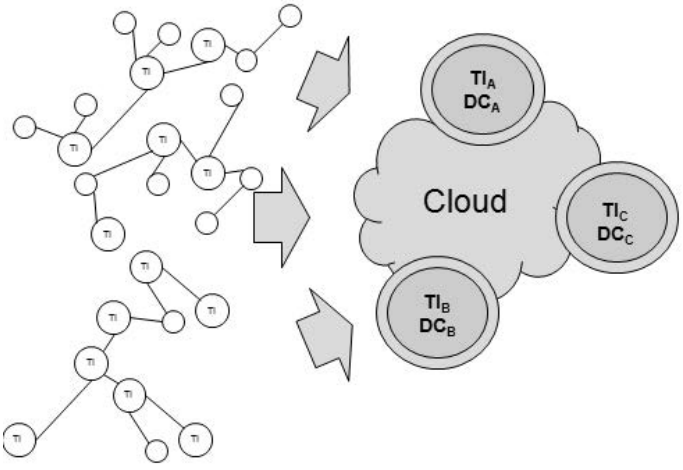


Figura 2-1 Mudança Organizacional alicerçada pela TI

A nova organização, agora, é uma combinação de diversas organizações, composta por células interconectadas com diversos pontos de acesso propiciados pela infraestrutura de TI. Com a popularização da INTERNET e a redução dos custos de conexão, a rede passou a ter uma abrangência maior. São tantos pontos de conexão que a figura da nuvem (CLOUD) para representar a rede parece ser mais adequada, conforme ilustra a **Figura 2-2**.

O elemento central de processamento e armazenamento dos dados e da informação na nuvem é o DATACENTER (DC), conforme ilustra a **Figura 2-2**. Ou seja, agora a TI é novamente centralizada em grandes pontos de armazenamento e processamento, os tais DATACENTERS, mas conserva a estrutura de interligação em redes. A nuvem, na verdade, é um conjunto de grandes pontos de armazenamento e processamento de informações conhecidos como DATACENTERS. Outras características citadas a seguir definem o que se convencionou chamar de CLOUD COMPUTING.


TI
TI
TI
TI
TI

A centralização ou consolidação em grandes estruturas como os DATACENTERS é viabilizada pela atual oferta de banda e pela existência de tecnologias que permitem alto poder de processamento e armazenamento em estruturas que reduzem o custo total de propriedade (TCO) da infraestrutura de TI.

É como se as organizações agora aproveitassem o melhor dos mundos. Centralizam o processamento e o armazenamento da informação em grandes estruturas, ao mesmo tempo em que estão integradas em rede.

Em resumo, a centralização da base de informação em estruturas como DATACENTERS e a implantação de políticas de infraestrutura, incluindo a segurança dos dados que permitam o fornecimento de níveis adequados de serviço para as aplicações, são o panorama que norteia a gestão da infraestrutura de TI.

CLOUD COMPUTING faz sentido? Segundo a IBM, 85% da capacidade de computação do mundo está ociosa. CLOUD COMPUTING é parte da evolução natural da TI e um meio para aprimorar o uso da capacidade computacional em todo o mundo.

O que seria a CLOUD COMPUTING? Na verdade, o conceito ainda se aprimora. A ideia inicial da CLOUD COMPUTING foi a de processar as aplicações e armazenar os dados fora do ambiente corporativo, otimizando o uso dos recursos. Este é o conceito atual de nuvem pública. A ideia central da CLOUD COMPUTING pública é permitir que as organizações executem boa parte dos serviços que hoje são executados em DATACENTERS corporativos ou em DATACENTERS providos por terceiros, saindo de um modelo baseado em Capex (custo de capital) para um modelo baseado em Opex (custo de operação) e onde agora os indicadores de desempenho estão atrelados aos níveis de serviço acordados entre clientes e fornecedores. Estes acordos idealmente deveriam variar em função da criticidade da aplicação, o que, na prática, é um desafio.

CLOUD COMPUTING é um passo adiante em relação a uma simples estrutura baseada em DATACENTER, quer seja público ou privado. A CLOUD COMPUTING significa mudar fundamentalmente a forma de operar a TI, saindo de um modelo baseado em aquisição de equipamentos para um modelo baseado em aquisição de serviços de TI. Alguns fornecedores continuam querendo vender hardware e, portanto, fomentam outras formas de nuvem.

Atualmente, os serviços de nuvem decorrentes da organização em nuvem pública são fornecidos, em sua grande maioria, por grandes organizações como Google, Microsoft, Amazon e outras organizações que hospedam e executam as aplicações dos clientes empresariais. A tendência é que esta oferta cresça e mais opções apareçam, aumentando a concorrência nesta área e reduzindo os custos de utilização destes serviços.

2.2. Conceito

CLOUD COMPUTING é um conjunto de recursos virtuais facilmente utilizáveis e acessíveis tais como hardware, software, plataformas de desenvolvimento e serviços. Estes recursos podem ser dinamicamente reconfigurados para se ajustarem a uma carga de trabalho (WORKLOAD) variável, permitindo a otimização do uso dos recursos. Este conjunto de recursos é tipicamente explorado através de um modelo pague-pelo-

-uso com garantias oferecidas pelo provedor através de acordos de nível de serviços (Vaquero et al, 2009).

CLOUD COMPUTING é substituir ativos de TI que precisam ser gerenciados internamente por funcionalidades e serviços do tipo pague conforme crescer a preços de mercado. Estas funcionalidades e serviços são desenvolvidos utilizando novas tecnologias como a VIRTUALIZAÇÃO, arquiteturas de aplicação e infraestrutura orientadas a serviço e tecnologias baseadas na Internet como meio de reduzir os custos de uso de recursos de hardware e software de TI usados para processamento, armazenamento e rede.

A **Figura 2-3** ilustra a mudança do modelo onde a capacidade do sistema acompanharia a demanda, promovendo a otimização do uso dos recursos. A **Figura 2-3 (a)** ilustra a situação comum onde a capacidade está sempre sobrando e a **Figura 2-3 (b)** ilustra a situação ideal.

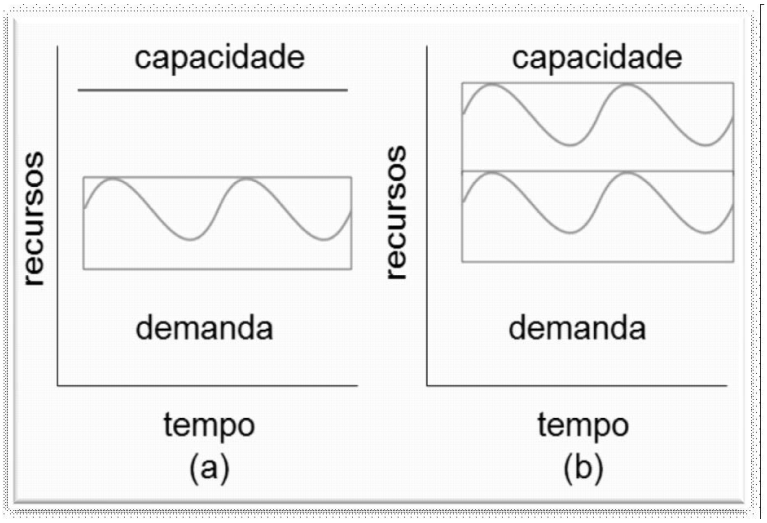


Figura 2-3 Mudança com CLOUD COMPUTING

O benefício da elasticidade da CLOUD COMPUTING, mostrado **na Figura 2-3 (b)**, permite transferir o risco da baixa utilização e da alta utilização (saturação) para uma situação de ajuste fino entre a carga de trabalho e os recursos disponíveis. A ideia central da CLOUD COMPUTING é possibilitar que as aplicações que rodam em DATACENTERS isolados possam rodar na nuvem (cuja tecnologia principal é a da INTERNET) em um ambiente de larga escala e de uso elástico de recursos.

Elasticidade tem a ver com a demanda atual e a demanda futura. Com o uso maciço da INTERNET, as organizações já não sabem exatamente como os clientes eletrônicos se comportam. Imagine uma companhia aérea vendendo bilhete a um preço irrisório no fim de semana e voltando a operação normal na segunda-feira. O que fazer com a infraestrutura de TI? No fim de semana precisa-se de muito recurso, mas já na segunda-feira os recursos estariam sobrando. Esta característica é difícil de ser obtida quando a organização utiliza uma infraestrutura interna. A VIRTUALIZAÇÃO utilizada internamente ajuda, mas pode não resolver. Como solicitar banda à operadora para o fim de semana

e depois devolver esta mesma banda de forma rápida? Existem questões ainda difíceis de serem resolvidas.

A proposta da CLOUD COMPUTING é de alguma forma melhorar o uso dos recursos e tornar a operação de TI mais econômica. Também é evidente que só a elasticidade propiciada pelos componentes da nuvem, os DATACENTERS, não é suficiente para garantir a elasticidade requerida pelas organizações no uso dos recursos. Também as aplicações e as suas arquiteturas deverão ser orientadas a serviço para garantir a elasticidade. Tradicionais "approaches" baseados em técnicas de escalabilidade horizontal (*scale-out*) e escalabilidade vertical (*scale-up*) utilizadas tanto para o hardware como para o software já não são mais suficientes.

2.3. Características Essenciais

Um modelo de CLOUD COMPUTING deve apresentar algumas características essenciais descritas abaixo:

- **Serviço sob demanda**: funcionalidades computacionais são providas automaticamente sem a interação humana com o provedor de serviço.
- **Amplo acesso aos serviços de rede**: recursos computacionais estão disponíveis através da INTERNET e são acessados via mecanismos padronizados, para que possam ser utilizados por dispositivos móveis e portáteis, computadores, etc.
- **Pooling de recursos**: recursos computacionais (físicos ou virtuais) do provedor são utilizados para servir a múltiplos usuários, sendo alocados e realocados dinamicamente conforme a demanda do usuário.
- **Elasticidade rápida**: as funcionalidades computacionais devem ser rápidas e elasticamente providas, assim como rapidamente liberadas. O usuário dos recursos deve ter a impressão de que ele possui recursos ilimitados, que podem ser adquiridos (comprados) em qualquer quantidade e a qualquer momento.
- **Medição de serviços**: os sistemas de gerenciamento utilizados pela CLOUD COMPUTING controlam e monitoram automaticamente os recursos para cada tipo de serviço (armazenamento, processamento e largura de banda). Esse monitoramento do uso dos recursos deve ser transparente para o provedor de serviços, assim como para o consumidor do serviço utilizado.

A proposta da CLOUD COMPUTING é criar a ilusão de que o recurso computacional é infinito e, ao mesmo tempo, permitir a eliminação do comprometimento antecipado da capacidade. Além disso, a ideia é permitir o pagamento pelo uso de recursos de forma granular. Do ponto de vista prático, os provedores, em sua grande maioria, ainda não estão preparados para disponibilizar esta forma de serviço.

2.4. Modelos de Serviços

Existem três principais modelos de serviços para CLOUD COMPUTING:

- **Infraestrutura como um Serviço** (*Infrastructure as a Service – IaaS*): capacidade que o provedor tem de oferecer uma infraestrutura de processamento e arma-

zenamento de forma transparente. Neste cenário, o usuário não tem o controle da infraestrutura física, mas, através de mecanismos de VIRTUALIZAÇÃO, possui controle sobre os sistemas operacionais, armazenamento, aplicações instaladas e possivelmente um controle limitado dos recursos de rede. Um exemplo de IaaS é a opção AMAZON EC2.

- **Software como um Serviço** (*Software as a Service – SaaS*): aplicações de interesse de uma grande quantidade de usuários passam a ser hospedadas na nuvem como uma alternativa ao processamento local. As aplicações são oferecidas como serviços por provedores e acessadas pelos clientes por aplicações como o browser. Todo o controle e gerenciamento da rede, sistemas operacionais, servidores e armazenamento é feito pelo provedor de serviço. O Google Apps e o SalesForce.com são exemplos de SaaS.
- **Plataforma como um Serviço** (*Platform as a Service – PaaS*): capacidade oferecida pelo provedor para o desenvolvedor de aplicações que serão executadas e disponibilizadas na nuvem. A plataforma na nuvem oferece um modelo de computação, armazenamento e comunicação para as aplicações. Exemplos de PaaS são a AppEngine do Google e o AZURE da Microsoft.

A **Figura 2-4** ilustra as possíveis arquiteturas para a CLOUD COMPUTING quando comparadas ao serviço de DATACENTER convencional conhecido como HOSTING.

Os conceitos de SaaS e IaaS já são bem conhecidos e, de alguma forma, já utilizados por diversas organizações. SaaS trata da troca de um modelo de software baseado em venda de licenças (Capex) por um modelo baseado no uso do software como serviço (Opex). O modelo IaaS trata da utilização da infraestrutura de terceiros como serviço. Os requisitos de banda, latência, poder de processamento são alterados para garantias do nível de disponibilidade e de desempenho.

O conceito de PaaS, menos conhecido, está vinculado ao uso de ferramentas de desenvolvimento de software gratuitas, oferecidas por provedores de serviços, onde os desenvolvedores criam as aplicações e as desenvolvem utilizando a Internet. Neste caso, o provedor oferece a plataforma de desenvolvimento.

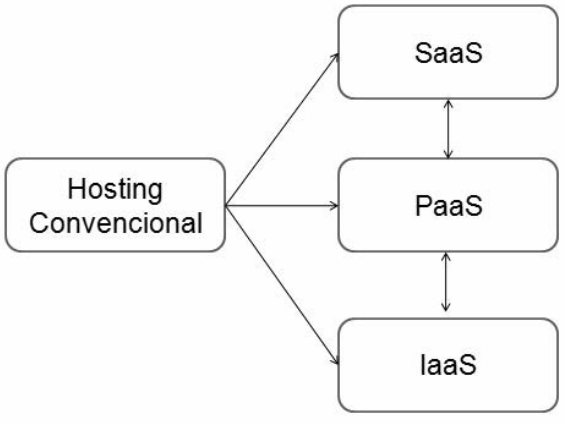


Figura 2-4 Arquiteturas de CLOUD COMPUTING

Para entender melhor a CLOUD COMPUTING, pode-se tentar identificar os papéis desempenhados na arquitetura baseada em nuvem. A **Figura 2-5** destaca quem fornece serviços e quem consome. IaaS suporta PaaS, que suporta SaaS.

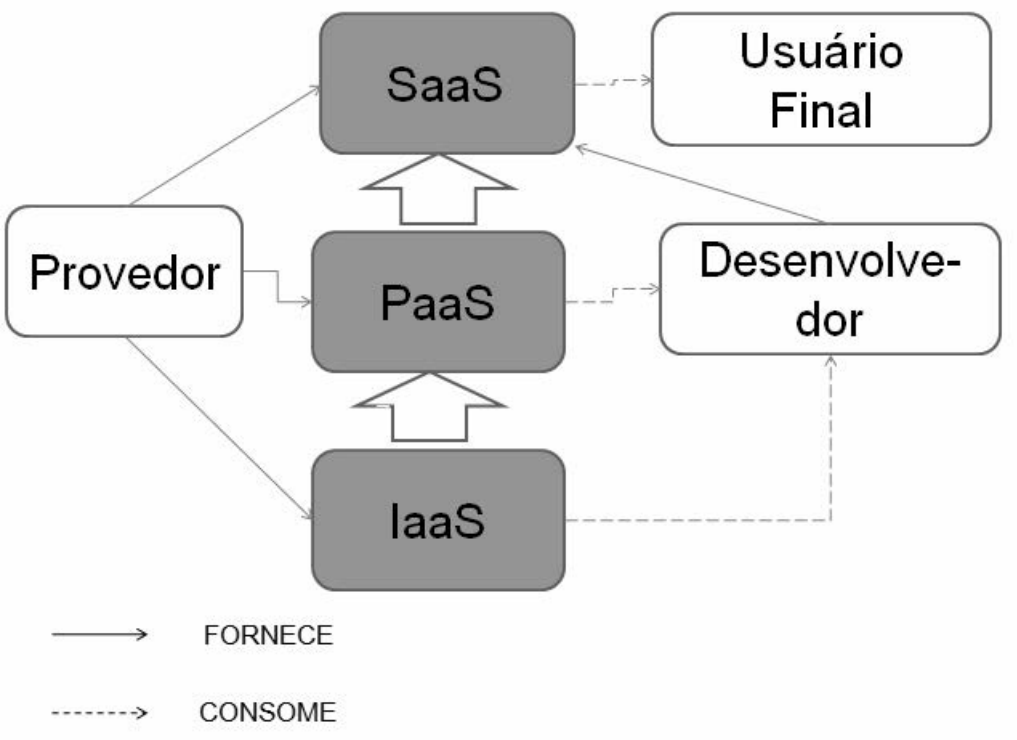


Figura 2-5 Papéis em CLOUD COMPUTING

O provedor de serviços é responsável por disponibilizar, gerenciar e monitorar toda a estrutura para a solução de CLOUD COMPUTING, deixando os desenvolvedores e usuários finais sem esses tipos de responsabilidade. Para isso, o provedor pode fornecer serviços nas três modalidades. Os desenvolvedores utilizam os recursos fornecidos e provêm serviços para os usuários finais.

Esta organização em papéis ajuda a definir os atores e os seus diferentes interesses em CLOUD COMPUTING. Os atores podem assumir vários papéis ao mesmo tempo de acordo com os interesses, sendo que apenas o provedor fornece suporte a todos os modelos de serviços.

Do ponto de vista de interação entre os três modelos de serviços, a IaaS fornece recursos computacionais, seja de hardware ou software, para a PaaS que, por sua vez, fornece recursos, tecnologias e ferramentas para o desenvolvimento e execução dos serviços implementados a serem disponibilizados como SaaS. Importante ressaltar que uma organização provedora de serviços de nuvem não precisa obrigatoriamente disponibilizar os três modelos. Ou seja, um provedor de serviços de nuvem pode disponibilizar a opção IaaS sem ter que obrigatoriamente disponibilizar também uma plataforma PaaS na nuvem.

2.5. Modelos de Implantação

Existem três principais modelos de implantação da CLOUD COMPUTING:

- **Nuvem privada** (*private cloud*): compreende uma infraestrutura de nuvem operada e quase sempre gerenciada pela organização cliente. Os serviços são oferecidos para serem utilizados internamente pela própria organização, não estando publicamente disponíveis para uso geral. Em alguns casos pode ser gerenciada por terceiros. Pesquisa realizada pela Forbes Insights[3] com 235 CIOs e outros executivos de grandes companhias americanas aponta que 52% das organizações estão avaliando a adoção de nuvens privadas.
- **Nuvem pública** (*public cloud*): nuvem é disponibilizada publicamente através do modelo pague-por-uso. São oferecidas por organizações públicas ou por grandes grupos industriais que possuem grande capacidade de processamento e armazenamento. A mesma pesquisa realizada pela Forbes Insights e citada anteriormente aponta que 38% das organizações estão avaliando também a adoção de nuvens públicas.
- **Núvem híbrida** (*hybrid cloud*): a infraestrutura é uma composição de duas ou mais nuvens (privadas, públicas ou comunidade) que continuam a ser entidades únicas, porém conectadas através de tecnologias proprietárias ou padronizadas que propiciam a portabilidade de dados e aplicações.

A **Figura 2-6** ilustra os modelos de nuvem.

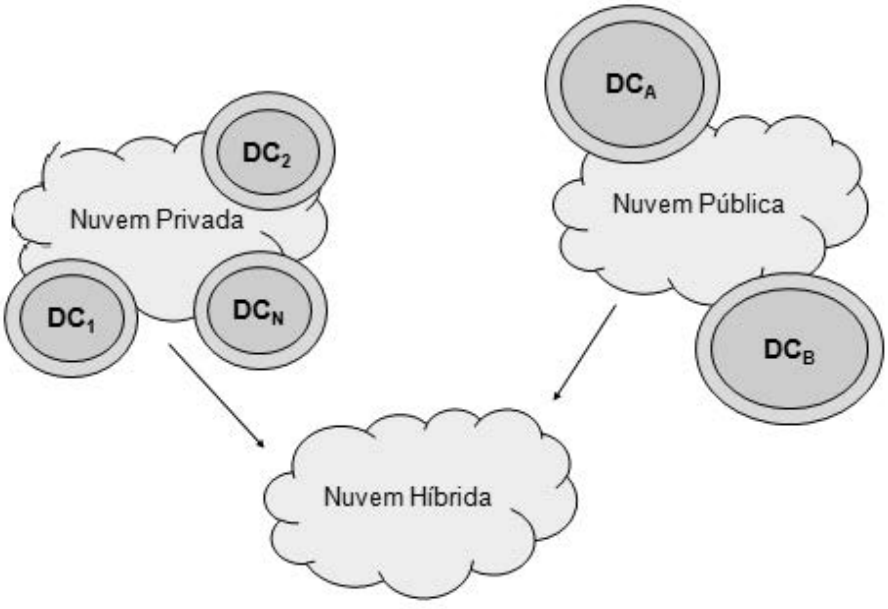


Figura 2-6 Nuvem Híbrida

3 Seeding the Cloud: Enterprises Set Their Strategies for Cloud Computing, Forbes Insights, 2010.

2.6. Benefícios

Os possíveis benefícios do uso de um modelo de CLOUD COMPUTING por organizações são:

- **Menores custos de infraestrutura**: a ideia de pagar pelo uso e eliminação dos custos de capital induz a obter menores custos com infraestrutura.
- **Aumento da utilização da infraestrutura**: a infraestrutura compartilhada e sob demanda passa a ser mais utilizada e permite a redução do custo por parte do provedor.
- **Aumento da segurança**: uma infraestrutura centralizada pode melhorar a segurança incluindo rotinas de backup otimizadas e testadas. (Há controvérsias.)
- **Acesso a aplicações sofisticadas**: aplicações consideradas caras podem ser utilizadas no modelo sob demanda.
- **Economia de energia**: redução de custos de energia e refrigeração é possível em uma estrutura centralizada como a proposta pela nuvem.
- **Aumento da produtividade por usuário**: usuários podem utilizar as aplicações de qualquer lugar, o que pode impactar positivamente a produtividade.
- **Aumento da confiabilidade**: existência de estrutura de contingência quase que obrigatória em CLOUD COMPUTING pode melhorar a confiabilidadade das aplicações.
- **Escalabilidade sob demanda**: a organização não precisa projetar a infraestrutura pelo pico. A computação em nuvem permite alocar recursos sob demanda, tornando a escalabilidade uma realidade.

2.7. Desafios

Existem diversos desafios para a adoção da CLOUD COMPUTING. Os principais seriam:

- **Falta de interoperabilidade**: a maioria dos modelos adotados pelos fornecedores de sistemas de CLOUD COMPUTING é integrada verticalmente e limita a escolha da plataforma.
- **Compatibilidade entre aplicações**: aplicações construídas para a CLOUD COMPUTING são muitas vezes incompatíveis com as aplicações legadas.
- **Dificuldades em obedecer a normas regulatórias**: a regulação pode limitar o uso da CLOUD COMPUTING para alguns ambientes. Nuvens internas podem ser uma opção mais adequada para diversas organizações.
- **Segurança inadequada**: em muitas situações a segurança pode ser o gargalo da CLOUD COMPUTING. Os provedores de CLOUD COMPUTING suportam arquitetura *multitenant*, na qual todos os usuários e aplicativos compartilham uma única infraestrutura e base de códigos mantida centralmente. Em comparação com as aplicações cliente/servidor, os clientes em aplicações *multitenant* compartilham

a mesma instância física e versão de um aplicativo. As implantações individuais desses aplicativos ocupam partições virtuais, em vez de pilhas físicas separadas de hardware e software. Esta arquitetura, mesmo sendo interessante para diversas aplicações, pode trazer problemas com a segurança.

2.8. Iniciativas

Diversas iniciativas surgem baseadas na CLOUD COMPUTING. Serão tratadas aqui as principais iniciativas mundiais nesta área. A **Figura 2-7** ilustra o posicionamento dos principais atores tecnológicos envolvidos em fornecer serviços de nuvem privada. O modelo do Google, Amazon e Microsoft é baseado no fornecimento de serviços de nuvem por fornecedor exclusivo. Os modelos da HP, ORACLE e da IBM continuam focados na verticalização da oferta que vai desde os equipamentos de hardware, passa pelo sistema operacional e vai até as aplicações. O modelo da VMware, Cisco e EMC (VCE) passa pela combinação do melhor dos mundos dos três fabricantes, mas deixa as camadas de cima da pilha de TI para outros fornecedores.

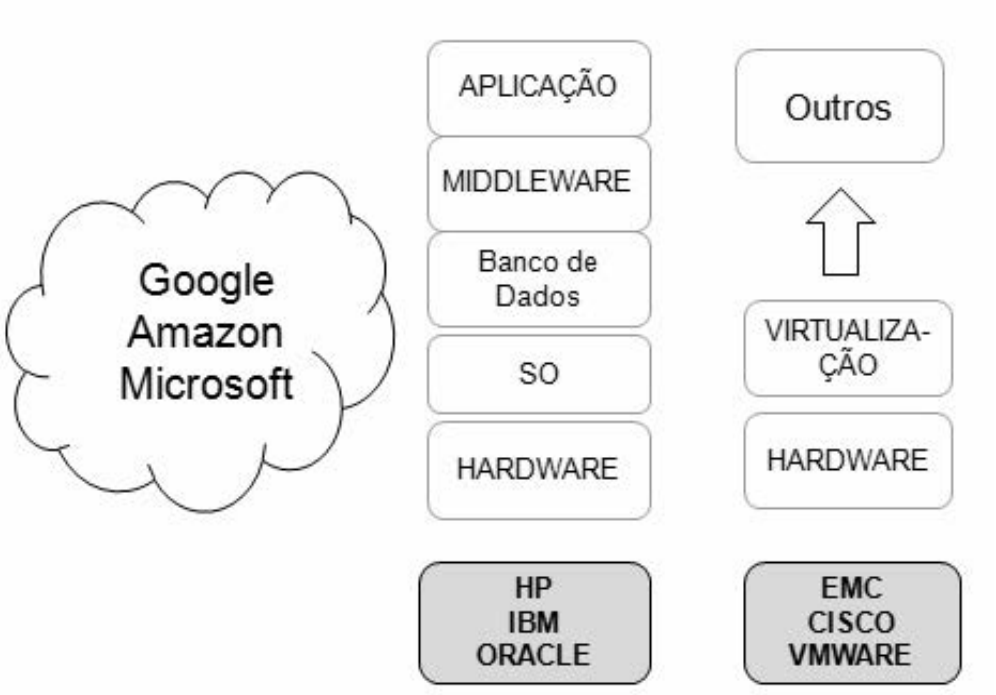


Figura 2-7 Iniciativas dos Fornecedores de Nuvem

2.8.1. Conceito na Prática: VMware vSphere

A VMware disponibiliza o *VMware vSphere Cloud,* um sistema operacional para CLOUD COMPUTING. A VMware sugere que o *vSphere Cloud* permite às organizações tranformarem seus DATACENTERS internos em uma nuvem pública, privada ou híbrida. A **Figura 2-8** ilustra as três opções.

Nuvem Privada

O *vSphere Cloud* permite que as organizações transformem seus DATACENTERS em nuvens privadas que podem ser mais eficientes no fornecimento das aplicações existentes. A nuvem privada permite a alocação de recursos de forma dinâmica e automação centralizada. Adotando o vSphere como uma plataforma privada de nuvem, as equipes de TI podem prestar serviços de manutenção de uma forma flexível e mais eficiente ao automatizar os processos chaves, tais como a gestão dos sistemas e o provisionamento da aplicação. A nuvem privada é um conceito adotado também por grandes fabricantes como a EMC.

Nuvem Pública

A ideia de nuvem pública da VMware é propiciar uma plataforma que assegure a otimização do uso dos recursos junto com a elasticidade para fornecer os níveis de serviço adequados para as aplicações. Com um grande ecossistema de fornecedores de serviço de nuvem, as organizações podem começar a utilizar os serviços de virtualização *on-demand*, pagar-pelo-uso e utilizar serviços prontos de infraestrutura pública.

Nuvem Híbrida

A visão de VMware para CLOUD COMPUTING é construir uma ponte entre recursos internos e externos disponíveis, permitindo que os clientes criem uma infraestrutura híbirida de nuvem. O modelo permite que as organizações consigam obter benefícios reais da nuvem, que seriam uma mistura estratégica de soluções internas e externas para o armazenamento de dados, a hospedagem de aplicações e a recuperação de desastres. Segundo a VMware, os clientes que adotam soluções privadas ou públicas de nuvem baseadas no vSphere estariam em melhor posição neste momento se utilizassem um modelo híbrido de nuvem.

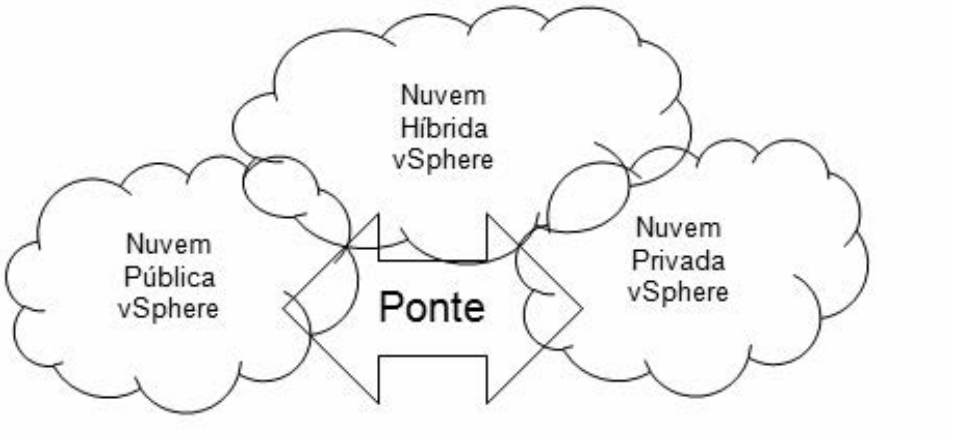


Figura 2-8 VMware vCloud

vCloud Express

A VMware oferece um serviço do tipo IaaS, o vCLoud Express, para provedores parceiros do tipo ISV (*independent software vendor* – fornecedor independente de software) e a comunidade *open-source,* que pode se integrar ao ambiente vCloud através de uma API chamada de vCLoud . O vCloud Express é um serviço *on-demand*, pague-pelo-uso, compatível com o ambiente VMware e com os serviços virtualizados baseados no VMware.

Esta classe de serviço permite reduzir os desafios do alto custo de capital para utilização de recursos de infraestrutura cuja demanda varia ao longo do tempo.

2.8.2. Conceito na Prática: Windows AZURE

Aplicações locais sempre são construídas em algum tipo de plataforma que inclui um sistema operacional, uma forma de armazenamento de dados e outros componentes. As aplicações executadas na nuvem precisam de uma plataforma similar.

O objetivo do Microsoft Windows AZURE é fornecer essa plataforma que pode ser usada para a execução de aplicações Windows e o armazenamento de dados na nuvem. A **Figura 2-9** ilustra a ideia central do Windows AZURE. *Desktops* de empresas e de consumidores se relacionam com aplicações e dados armazenados em servidores colocados na nuvem. O relacionamento agora passa a ser ditado por aquisição de serviços e não por aquisição de licença de software.

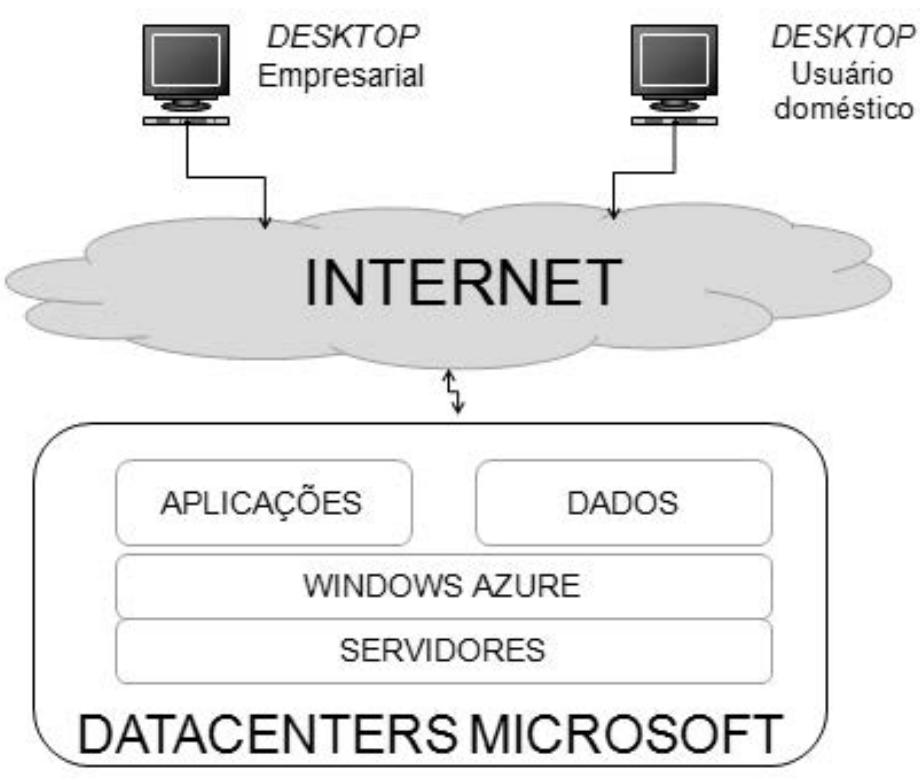


Figura 2-9 Windows AZURE

A **Figura 2-10** ilustra um típico DATACENTER da Microsoft, componente central da nuvem Microsoft.

Figura 2-10 Microsoft Dublin DATACENTER

O Windows AZURE é executado em máquinas presentes nos DATACENTERS da Microsoft. Em vez de fornecer o software para que os clientes instalem e executem as aplicações em seus próprios computadores, o Windows AZURE é um serviço: os clientes o utilizam para executar aplicações e armazenar dados em máquinas que pertencem à Microsoft. Essas aplicações podem fornecer serviços para empresas, consumidores ou ambos. Alguns exemplos de aplicações que podem ser construídas no Windows AZURE são:

- ISV (*Independent Software Vendor*) pode criar uma aplicação para usuários corporativos utilizando SaaS (*Software as a Service* – Software como Serviço). Os ISVs podem usar o Windows AZURE como base para várias aplicações SaaS orientadas aos negócios.
- ISV pode criar uma aplicação SaaS voltada para os consumidores. O Windows AZURE foi projetado para dar suporte a softwares escalonáveis, portanto uma empresa que pretenda atingir um grande mercado de consumidores pode escolhê-lo como base para uma nova aplicação.
- As empresas podem usar o Windows AZURE para construir e executar aplicações que serão usadas pelos próprios funcionários. Embora essa situação não exija a escalabilidade de uma aplicação voltada para os consumidores, a confiabilidade e o gerenciamento oferecidos pelo Windows AZURE podem fazer dela uma opção bastante interessante.

Seja qual for a finalidade da aplicação construída no Windows AZURE, a plataforma oferece os mesmos componentes básicos, como mostra a **Figura 2-11**.

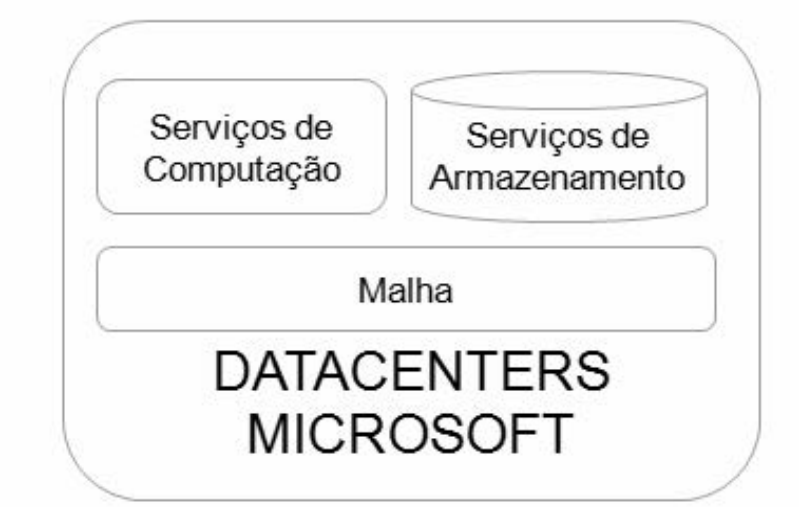


Figura 2-11 Componentes do Windows AZURE

Como os próprios nomes sugerem, o serviço de computação executa aplicações e o serviço de armazenamento armazena os dados. O terceiro componente, a malha do Windows AZURE, oferece uma maneira de gerenciar e monitorar as aplicações que usam essa plataforma na nuvem.

Serviço de Computação

O serviço de computação do Windows AZURE pode executar vários tipos de aplicações. O principal objetivo da plataforma, no entanto, é dar suporte a aplicações que têm um grande número de usuários simultâneos. Alcançar esse objetivo escalando verticalmente, ou seja, executando aplicações em máquinas cada vez maiores, não é possível. Em vez disso, o Windows AZURE foi projetado para dar suporte a aplicações que escalam horizontalmente (*scale-out*), executando múltiplas cópias do mesmo código em vários servidores padrão da indústria.

Para isso, a aplicação Windows AZURE pode ter múltiplas instâncias, cada uma executada em sua própria VM (*virtual machine* – máquina virtual). Essas VMs são executadas no Windows Server 2008 de 64 bits e fornecidas por um HYPERVISOR (baseado no Hyper-V da Microsoft) que foi modificado para ser usado na nuvem da Microsoft. Para executar uma aplicação, o desenvolvedor acessa o portal do Windows AZURE em seu navegador da Web utilizando como identificador o Windows Live ID. Depois escolhe se quer criar uma conta de hospedagem para executar as aplicações, uma conta de armazenamento para armazenar dados, ou ambas. Se escolher a conta de hospedagem, poderá então fazer um upload de sua aplicação, especificando o número de instâncias exigido por ela. O Windows AZURE cria então as VMs necessárias e executa a aplicação.

É importante notar que o desenvolvedor não pode fornecer sua própria imagem de VM para ser executada pelo Windows AZURE. A plataforma fornece e mantém sua própria cópia do Windows. Os desenvolvedores se concentram apenas na criação das aplicações que serão executadas no Windows AZURE.

Serviço de Armazenamento

As aplicações trabalham com dados de várias formas diferentes. Por isso, o serviço de Armazenamento do Windows AZURE oferece várias opções. O modo mais simples de armazenar dados no Windows AZURE é através dos *blobs*. Um *blob* contém dados binários e uma hierarquia simples: a conta de armazenamento pode ter um ou mais contêineres, e cada um deles possui um ou mais *blobs*. Os *blobs* podem ser grandes – até 50 GB cada – e também podem ter metadados associados, como informações sobre onde foi tirada uma fotografia JPEG ou quem é o cantor de um arquivo MP3.

Os *blobs* são apropriados para algumas situações, mas para outras são muito desestruturados. Para que as aplicações trabalhem com os dados de maneira mais refinada, o armazenamento do Windows AZURE oferece tabelas. Essas tabelas não são relacionais. Na verdade, embora sejam chamadas de "tabelas", seus dados são armazenados em uma hierarquia simples de entidades que contêm propriedades. E em vez de usar o SQL, a aplicação acessa os dados da tabela usando convenções definidas. A razão para o uso desse método é que ele possibilita um armazenamento com escalabilidade horizontal, ou seja, distribui os dados entre várias máquinas – com muito mais eficiência que um banco de dados relacional padrão. Na verdade, uma única tabela do Windows AZURE pode conter bilhões de entidades com terabytes de dados.

A principal função dos *blobs* e das tabelas é armazenar e acessar dados. A terceira opção no armazenamento do Windows AZURE, as filas, tem um propósito bastante diferente. Por meio das filas, as instâncias da função Web podem se comunicar com as instâncias da função "Trabalhador", função definida no próprio AZURE. Por exemplo, o usuário pode enviar uma solicitação para realizar uma tarefa com uso intensivo de computação através de uma página Web implementada por uma função Web do Windows AZURE. A instância da função Web que receber a solicitação poderá gravar uma mensagem em uma fila, descrevendo o trabalho a ser feito. A instância da função "Trabalhador" que estiver servindo essa fila poderá então ler a mensagem e desempenhar a tarefa especificada. Os resultados podem ser retornados através de outra fila ou manipulados de outra maneira.

Malha

Todas as aplicações Windows AZURE e todos os dados do armazenamento do Windows AZURE residem em DATACENTERS da Microsoft. Dentro destes DATACENTERS o conjunto de máquinas dedicadas ao Windows AZURE é organizado em uma malha.

A malha do Windows AZURE consiste em um grupo (grande) de máquinas gerenciadas por um software chamado controlador de malha. O controlador de malha é replicado em um grupo de cinco a sete máquinas, e todos os recursos da malha pertencem a ele: computadores, switches, balanceadores de carga, entre outros. Ele pode se comunicar com um agente de malha em cada computador, portanto ele reconhece todas as aplicações Windows AZURE presentes nessa malha. (É interessante notar que o controlador de malha vê o armazenamento do Windows AZURE apenas como mais uma aplicação, portanto os detalhes da replicação e do gerenciamento dos dados não ficam visíveis para o controlador).

2.8.3. Conceito na Prática: Amazon AWS

Os serviços AWS fornecidos pela Amazon são o que há de mais próximo do modelo teórico proposto para a CLOUD COMPUTING. A Amazon vem desenvolvendo estes serviços de nuvem há alguns anos e hoje possui uma oferta única de serviços de nuvem.

Os serviços AWS permitem o acesso a recursos de computação, armazenamento e banco de dados e outros serviços de infraestrutura *on demand*. A ideia é que esta forma de computação reduza custos, melhore o fluxo de caixa da organização contratante, minimize os riscos do negócio e maximize as oportunidades. Segundo a Amazon, o serviço tradicional de DATACENTER, quando comparado ao serviço de nuvem, é pouco eficaz considerando que o uso e a capacidade dos recursos estão otimizados no modelo AWS. A **Figura 2-12** ilustra os principais serviços AWS da Amazon.

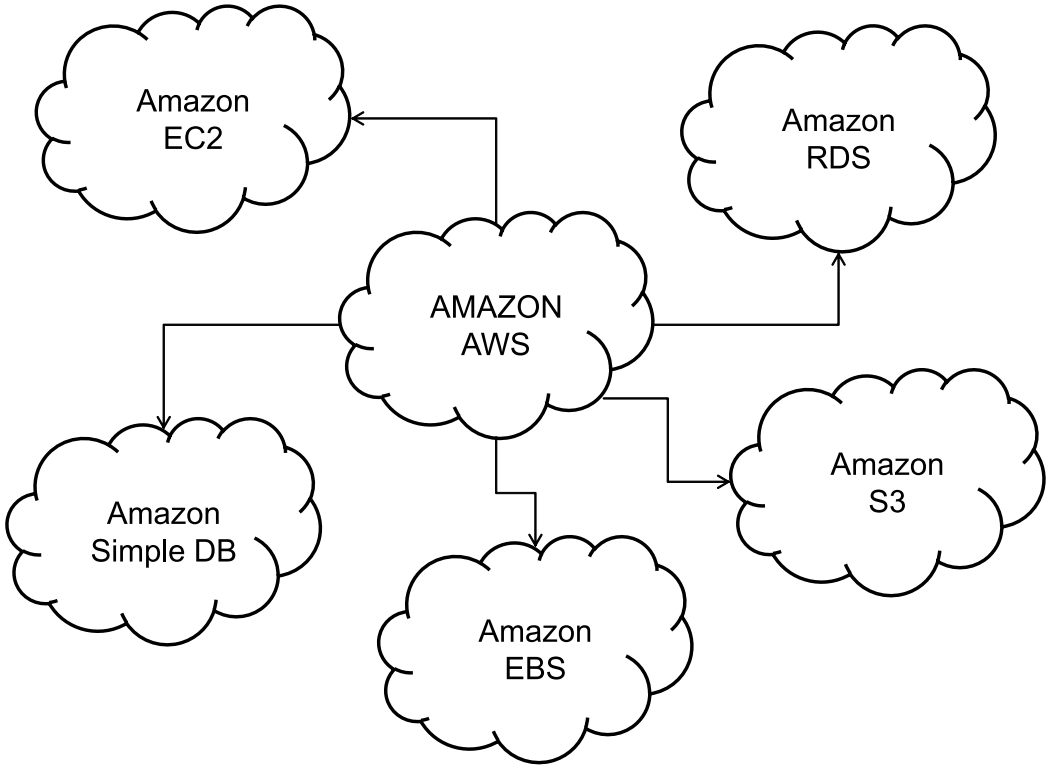


Figura 2-12 Serviços AWS da AMAZON

Serviços AWS

- ***Amazon Elastic Compute Cloud (Amazon EC2):*** o Amazon EC2 é um serviço web que permite aos desenvolvedores utilizar os recursos de nuvem sob demanda. O Amazon EC2 permite pagar pelo uso. A criação de novas instâncias de servidor pode ser realizada em minutos. Também permite que as aplicações construídas estejam isoladas e livres de cenários de falha comuns em ambientes não controlados. A opção *Auto Scaling* permite que a capacidade EC2 cresça ou diminua de acordo com a demanda, minimizando os custos. O *Auto Scaling* é útil para

aplicações que variam o uso semanalmente ou até diariamente. O *Auto Scaling* é disponibilizado pelo serviço *Amazon CloudWatch*.

- ***Amazon SimpleDB:*** O Amazon SimpleDB é um armazenamento de dados do tipo não relacional altamente disponível, escalável e flexível que alivia o trabalho de administração de banco de dados. Os desenvolvedores simplesmente armazenam itens de dados de consulta via serviços web e o Amazon SimpleDB faz o resto. *Amazon SimpleDB* é otimizado para fornecer alta disponibilidade, facilidade de escalabilidade e flexibilidade, com pouca ou nenhuma carga administrativa. O *Amazon SimpleDB* cria e gerencia várias réplicas dos dados geograficamente distribuídos para habilitar a alta disponibilidade de dados e a sua confiabilidade. O serviço responde a alterações no tráfego cobrando apenas para recursos de computação e armazenamento efetivamente consumidos. Pode-se alterar seu modelo de dados em tempo real e os dados são indexados automaticamente. Com o *Amazon SimpleDB*, pode-se concentrar no desenvolvimento de aplicações sem se preocupar sobre provisionamento de infraestrutura, alta disponibilidade, manutenção de software, gerenciamento de esquema e índice ou ajuste de desempenho.
- ***Amazon Relational Database Service (Amazon RDS):*** o serviço de banco de dados relacional *Amazon RDS* é um serviço web que torna mais fácil configurar, operar e dimensionar um banco de dados relacional na nuvem. O serviço fornece capacidade a um custo adequado e é redimensionável enquanto gerencia tarefas de administração de banco de dados normalmente demoradas, liberando o analista para se concentrar em seus aplicativos e negócios. O Amazon RDS permite acesso a todos os recursos do banco de dados MySQL. Isso significa que o código, as aplicações e as ferramentas que o usuário já usa hoje em bancos de dados MySQL, podem continuar existindo no *Amazon RDS*. O Amazon RDS aplica automaticamente patches do software de banco de dados e faz o backup do banco de dados, armazenando-os por um período de retenção definido pelo usuário. No RDS pode-se beneficiar da flexibilidade de poder dimensionar os recursos de computação ou capacidade de armazenamento associado com a instância do banco de dados relacional através de uma única chamada de API. Além disso, o *Amazon RDS* permite implantar a instância de banco de dados em várias zonas de disponibilidade para conseguir melhorar a disponibilidade e confiabilidade para implantações de produção críticos.
- ***Amazon Simple Storage Service (Amazon S3):*** O *Amazon S3* é um armazenamento para a Internet. Ele foi projetado para fazer a utilização da web mais fácil para desenvolvedores. O *Amazon S3* fornece uma interface de serviços web simples que pode ser usada para armazenar e recuperar qualquer quantidade de dados, a qualquer momento, de qualquer lugar na web, segundo a Amazon. O S3 permite que o desenvolvedor utilize uma infraestrutura escalável, confiável, segura, rápida e barata que a Amazon mesmo usa para executar sua própria rede global de sites da web. O serviço visa maximizar os benefícios da escala e passar esses benefícios para os desenvolvedores.

- ***Amazon Elastic Block Store (Amazon EBS):*** o *Amazon Elastic Block Store* (EBS) fornece volumes de armazenamento no nível de bloco para uso com instâncias Amazon EC2. Os volumes *Amazon EBS* são armazenados e sua persistência é independente da vida de uma instância. O *Amazon EBS* fornece volumes de armazenamento altamente disponíveis e altamente confiáveis que podem ser anexados a uma instância em execução do **Amazon EC2** e expostos como um dispositivo dentro da instância. Este recurso é particularmente adequado para aplicativos que exigem um banco de dados, sistema de arquivos ou acesso ao armazenamento no nível de bloco bruto.

Centros de Suporte dos Serviços AWS

A Amazon possui dois centros com foco em aspectos de custo e segurança. O foco destes centros é o de ajudar os clientes a entenderem as reais vantagens do ambiente de nuvem em termos de custo, benefício e segurança.

- ***AWS Economics Center:*** O *AWS Economics Center* fornece acesso a informações, ferramentas e recursos para comparar os custos da Amazon Web Services com alternativas convencionais de infraestrutura de TI. O objetivo é ajudar os desenvolvedores e líderes empresariais a quantificar os benefícios econômicos da CLOUD COMPUTING.
- ***AWS Security Center:*** O *AWS Security Center* fornece links para informações técnicas, ferramentas e orientação para ajudar a criar e gerenciar aplicativos seguros na nuvem AWS. O objetivo é usar este fórum para notificar proativamente os desenvolvedores sobre boletins de segurança.

2.9. Riscos e Segurança

Existem várias características importantes no modelo de CLOUD COMPUTING e a combinação destas características fornecidas pelo modelo adotado pelos provedores fazem o risco variar.

O Gartner recentemente sugeriu alguns cuidados que o cliente deve ter para mitigar o risco referente à aquisição de serviços de um provedor de CLOUD COMPUTING, descrito a seguir:

- Saber como é feito o acesso dos usuários;
- Saber como o provedor obedece às normas de regulação;
- Saber onde se localizam dos dados;
- Saber como os dados são segregados;
- Saber como os dados são recuperados;
- Saber como é feito o suporte;
- Entender a viabilidade do provedor no longo prazo.

O aspecto chave a ser avaliado na opção de entregar os serviços de TI para a CLOUD COMPUTING é o risco da perda do controle dos dados internos – e ele de fato existe.

Mesmo aplicações que eventualmente continuem a rodar localmente podem utilizar serviços de infraestrutura providos pela nuvem como armazenamento de dados, por exemplo, e funcionar utilizando serviços providos internamente e externamente. Normalmente os riscos aumentam até nestes casos.

O que deve ficar claro é que o ambiente de CLOUD COMPUTING é essencialmente diferente do ambiente tradicional de computação. Muda-se de um modelo amparado em equipamentos para um modelo orientado a serviços. Os acordos de nível serviço (*service level agreement – SLA*) passam a ser a interface natural entre provedores e organizações clientes.

A computação baseada em CLOUD COMPUTING pública adiciona novos atores à equação de segurança – provedores de serviço terceirizados; fornecedores de infraestrutura e funcionários diminuindo o controle que a TI exerce sobre essas três áreas.

As ofertas de serviços de nuvem ainda são confusas e, na medida em que estas ofertas crescem, os aspectos de segurança precisam ser rapidamente considerados. A oferta na forma de nuvem é uma opção ao DATACENTER corporativo e é preciso considerar os novos atores nesta nova forma de adquirir serviços de infraestrutura de TI. A confiança é o aspecto central na ida para a nuvem.

Os serviços de nuvem que acontecem nos formatos SaaS, PaaS e IaaS precisam ser elásticos, variando de acordo com as necessidades dos clientes, precisam ter níveis de segurança adequados à realidade dos clientes e das suas aplicações e precisam ser oferecidos em diversos locais. Também são ofertados para vários clientes que usam a infraestrutura de nuvem ao mesmo tempo otimizando o uso de recursos, mas tornando o aspecto de segurança complexo. A **Figura 2-13** ilustra essas características.

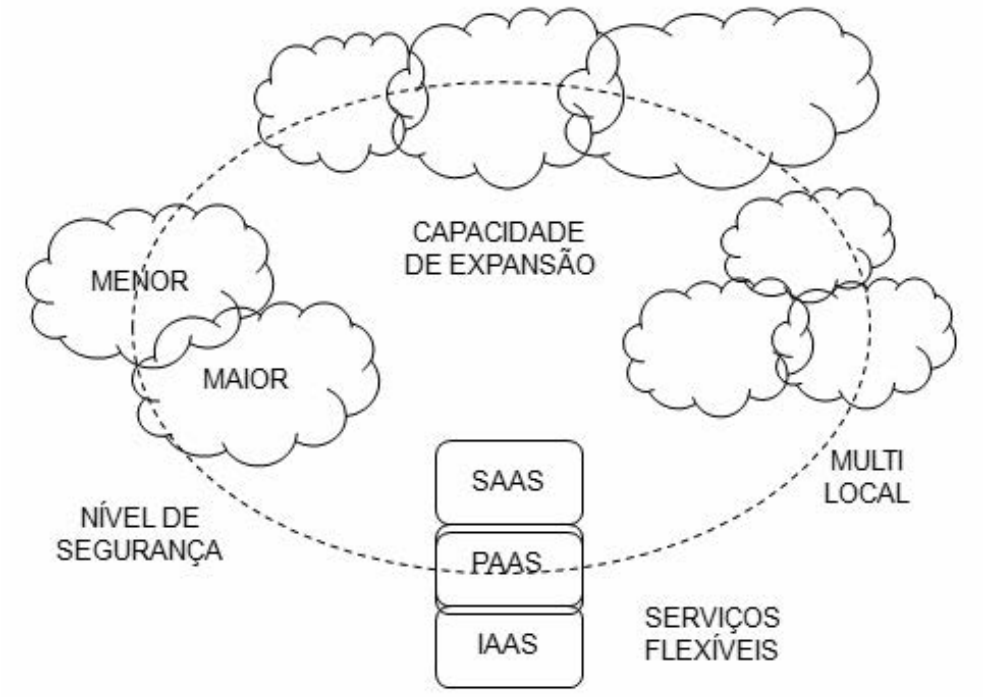


Figura 2-13 Recursos Essenciais de Nuvem (Fonte: RSA)

As aplicações que estão indo para a nuvem ainda são tímidas e de pouca confidencialidade, mas com o aumento exponencial da base de informação digital as empresas precisam considerar o fato de que respondem por boa parte destas informações em termos de segurança, privacidade e confidencialidade. O crescimento de uso da nuvem só será possível com o aumento da segurança. As empresas precisam assegurar a confidencialidade, integridade e disponibilidade dos dados no momento em que eles forem transmitidos, armazenados ou processados por terceiros na cadeia de serviços em nuvem.

Aspectos de segurança da infraestrutura tradicional mudam com a adoção da nuvem. As relações de confiança na cadeia da nuvem precisam ser cuidadosamente consideradas. O DATACENTER da nuvem agora é controlado por terceiros e estruturas convencionais de segurança já não servem mais. O controle passa a ser compartilhado e questões de responsabilidade vêm à tona: como saber quem teve acesso a que informações e aplicações? O acesso precisa ser individualizado. Diversas organizações trabalham mutualmente para definir os padrões de segurança na nuvem. A abrangência vai desde autenticação, autorização delegada, gerenciamento de chaves públicas, proteções contra perdas de dados até emissão de relatórios normativos.

A portabilidade na nuvem é outra questão a ser resolvida. Serão necessários arranjos que permitam portabilidade de identidade e informações entre várias nuvens. As questões de confidencialidade e privacidade são também muito relevantes. Abordagens de controle de acesso usuais em operações convencionais podem não funcionar na nuvem. Controle de acesso baseado em conteúdo e não na identidade do usuário é uma questão ainda não resolvida. O controle sobre informações confidenciais é uma questão muito relevante também. As empresas precisam saber se os provedores de serviço atendem a sua política de conformidade. Correlação de eventos internos e externos é outra demanda. Quem é o responsável pela violação dos dados?

Os níveis de serviço de segurança dos dados também deverão variar e devem atender a diferentes necessidades de confidencialidade, otimizando o TCO. A segurança deverá ser mantida durante todo o ciclo de vida dos dados. Os elementos principais para a proteção na nuvem são a identidade, a infraestrutura e as informações seguras. As empresas precisam ter confiança nos agentes da cadeia para abrir mão de algum tipo de controle.

- **Segurança de identidade**: o gerenciamento completo de identidades, os serviços de autenticação de terceiros e a identidade federada se tornarão elementos fundamentais para a segurança da nuvem. A segurança da identidade preserva a integridade e a confidencialidade dos dados e dos aplicativos, enquanto deixa o acesso prontamente disponível para os usuários apropriados. O suporte a esses recursos de gerenciamento de identidade para usuários e componentes da infraestrutura será um dos principais requisitos da computação em nuvem e a identidade precisará ser gerenciada de maneira que gere confiança.
- **Segurança das informações**: no DATACENTER tradicional, os controles sobre o acesso físico, o acesso a hardware e software e os controles de identidade se combinam para proteger os dados. Na nuvem, a barreira protetora que protege a

infraestrutura é diluída. Para compensar, a segurança passará a ser centrada nas informações. Os dados precisam de segurança própria que os acompanhe e os proteja.

- **Segurança da infraestrutura:** a infraestrutura de base da nuvem deve ser inerentemente segura, independentemente da nuvem ser privada ou pública ou prover serviço SAAS, PAAS ou IAAS.

A **Figura 2-14** ilustra os princípios para proteger a nuvem.

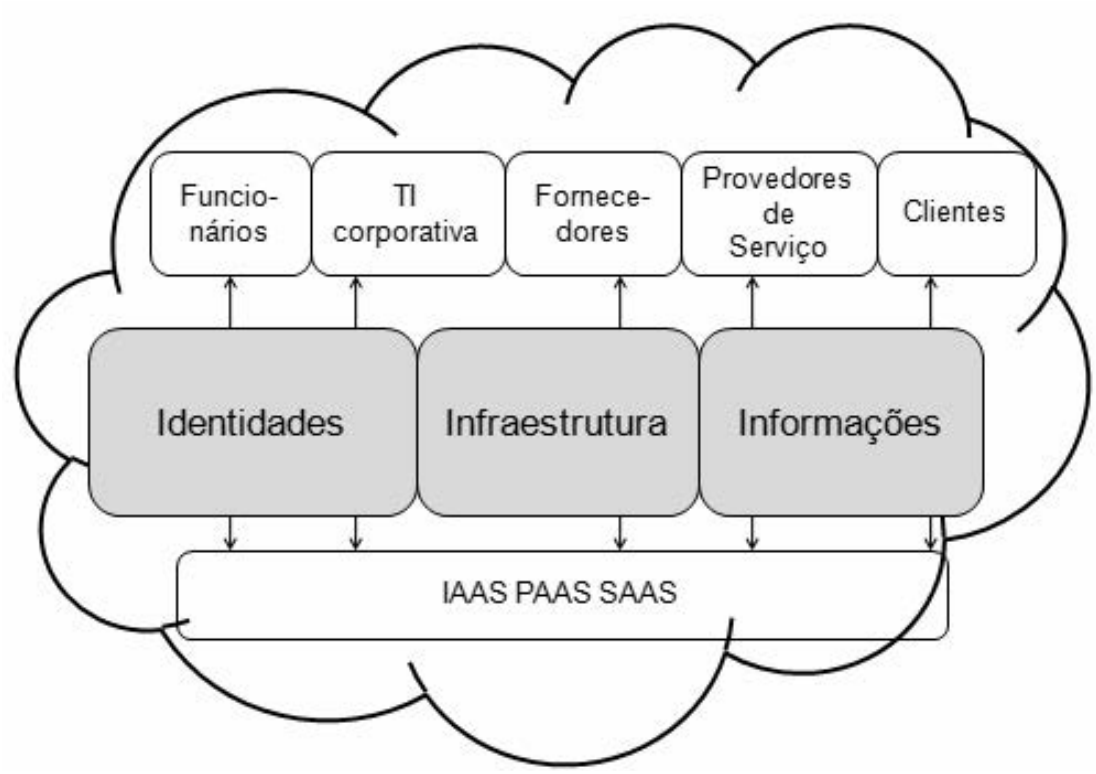


Figura 2-14 Princípios para proteger a Nuvem (Fonte: RSA)

A VMware e a Savvis sugerem algumas medidas operacionais de segurança interessantes que são resumidas aqui:

Do lado do provedor:

- Manter as redes que são partes da infraestrutura virtualizada isoladas. O isolamento pode ser feito através de segmentação com os switches ou utilizando VLANS. Outro método pode utilizar uma combinação destes dois métodos e o conceito de *virtual switch*.
- Manter as redes de gerenciamento isoladas. Manter as redes de gerenciamento isoladas e interfaces devidamente controladas.
- Manter as redes utilizadas para migração de máquinas virtuais e IP isoladas em redes não roteáveis. Estas redes precisam ser rápidas e ao mesmo tempo são suscetíveis a ataques.

- Manter as redes dos clientes isoladas. Estas redes devem ser isoladas das redes de gerenciamento e devem existir *firewalls* entre as redes para evitar problemas de segurança.
- Fornecer acesso seguro aos clientes de recursos de nuvem. Os clientes normalmente precisam ter acesso a recursos dentro da nuvem e para isto é necessário que o provedor disponibilize um portal com encriptação para efeito de segurança.
- Garantir *backups* e *restores* seguros e consistentes. O provedor deve garantir este aspecto básico de realizar o *backup* dos recursos e garantir um processo rápido de *restore*.
- Mecanismos fortes de autenticação, autorização e auditoria. Os provedores devem permitir a autenticação segura, acesso só aos recursos disponíveis e providenciar recursos de auditoria.
- Utilizar *templates* seguros de configuração e *gold images* do sistema operacional e das aplicações. Estas medidas reduzem os problemas com configurações inadequadas. Também o provisionamento de máquinas virtuais deve obedecer a critérios preestabelecidos antes de entrar em produção.
- Gerenciamento de recursos para evitar ataques do tipo DoS.

Do lado do cliente:

- Seguir práticas de mercado para a segurança dos sistemas operacionais. Os administradores devem seguir as mesmas regras utilizadas no ambiente interno relativas à segurança dos sistemas operacionais, como manter os últimos patches de configuração de segurança e utilizar *gold images*.
- Encriptar dados críticos. Dados críticos de sistemas importantes devem ser mantidos encriptados para dificultar ainda mais os possíveis ataques à rede.

2.10. Impactos da VIRTUALIZAÇÃO na CLOUD COMPUTING

A CLOUD COMPUTING, para funcionar adequadamente, depende das tecnologias utilizadas no DATACENTER, incluindo a VIRTUALIZAÇÃO. O esforço de entregar serviços sob demanda depende de DATACENTERS que precisam se ajustar a esta nova modalidade de serviços. De nada adianta ter as melhores tecnologias na construção do DATACENTER se o modelo utilizado para gerenciamento e oferta de serviços de TI não é orientado pela demanda. Diversos provedores de INTERNET possuidores de grandes DATACENTERS ainda não conseguem entregar a banda, armazenamento e processamento na forma de serviços e usam a VIRTUALIZAÇÃO de forma tímida.

A VIRTUALIZAÇÃO é o elemento central da nuvem, na medida em que permite aperfeiçoar o uso dos recursos e viabilizar o modelo de computação sob demanda.

2.11. Referências Bibliográficas

Amazon Web Services, disponível em www.amazon.com/aws.

Apresentando o Windows AZURE, David Chappel Associates, março de 2009.

Buyya, Rajkumar; Yeo, Chee Shin, Venugopal, Srikumar, Broberg, James, Brandic, Ivona. Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality for Delivering Computing as 5th utility, 2009.

Cloud Computing: Business Benefits with Security, Governance and Assurance Perspectives. ISACA, An ISACA Emerging Technology White Paper, 2009.

Flávio R. C. Sousa, Leonardo O. Moreira e Javam C. Machado. Computação em Nuvem: Conceitos, Tecnologias, Aplicações e Desafios, 2009.

Greenberg, Albert; Hamilton, James; Maltz, David A.; Patel, Parveen. The Cost of Cloud: Research problems in DATACENTER Networks, 2009.

Hilley, David. Cloud Computing: A Taxonomy of Platform and Infrastructure-level Offerings, CERCS Technical Report, 2009.

Marinos, A. and Briscoe, G. Community Cloud Computing. CoRR, abs/0907.2485, 2009.

O papel da segurança na computação em nuvem confiável. RSA, 2009.

Saten John, Deliver Cloud Benefits Inside your Walls, Forrester, 2009.

Securing the Cloud: A Review of Cloud Computing, Security Implications and Best Practices. VMware & Savvis, 2009.

Stanoevska-Slabeva; Katarina, Woznoak, Thomas, Ristol; Santi Editors. Grid and Cloud Computing: A Business Perspective on Technology and Applications, Springer, 2010.

UC Berkeley Reliable Adaptive Distributed Systems Laboratory. Above the Clouds: A Berkeley View of Cloud Computing, US Berkeley Reliable Adaptative Distributed Systems Laboratory, Feb 2009.

Vaquero, Luis M & Rodero-Merino, Luis & Caceres, Juan Lindner,Maik. A break in the Clouds: Towards a Cloud Definition. ACM Sigcomm Computer Communication Review, Volume 39, number 1, Jan 2009.

Varia, Jinesh. Architecting for the Cloud: Best Practices, Amazon Web Services, Jan 2010.

Veras, Manoel. DATACENTER: Componente Central da infraestrutura de TI, Brasport, 2009.

Verdi, Fábio Luciano; Rothenberg, Christian; Pasquini, Rafael, Magalhães, Maurício Ferreira, Novas Arquiteturas de DATACENTER para Cloud Computing, 2010.

VMware vSphere: The best platform for Cloud Infrastructures, VMware, 2010.

3. Tier Datacenter e Virtualização

3.1. Introdução

Um DATACENTER (ou DATA CENTER) é um conjunto integrado de componentes de alta tecnologia que permitem fornecer serviços de infraestrutura de TI de valor agregado, tipicamente processamento e armazenamento de dados, em larga escala, para qualquer tipo de organização. A denominação TIER DATACENTER faz referência à classificação em camadas adotada pela norma TIA-942, que norteia este capítulo.

O elemento central da Infraestrutura de TI de qualquer organização é o DATACENTER, e toda organização, de alguma forma, possui um DATACENTER ou utiliza serviços de um DATACENTER. Os DATACENTERS são alimentados por energia (*utilities*), possuem instalações (*facilities*) e servem a carga de TI, conforme ilustra a **Figura 3-1**. Em um projeto de DATACENTER a carga de TI define as instalações e a energia necessária.

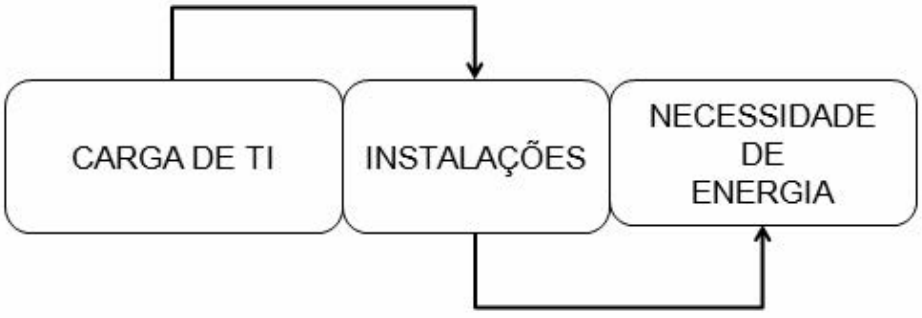


Figura 3-1 O DATACENTER

Os DATACENTERS e suas conexões formam a nuvem e podem fazer parte de arranjos de nuvem públicos e/ou privados. A grande confusão que se faz é achar que só existe DATACENTER nos provedores de serviço de INTERNET. Nos Estados Unidos (EUA), por exemplo, boa parte das organizações de médio e grande porte possui um DATACENTER próprio. O IDC apontou que nos EUA, em 2008, já existiam cerca de 2,5 milhões de DATACENTERS. Para efeito didático, os DATACENTERS podem ser divididos em cinco grandes blocos: instalações (*facilities*); energia; refrigeração; gerenciamento e a carga de TI, conforme ilustra a **Figura 3-2**.

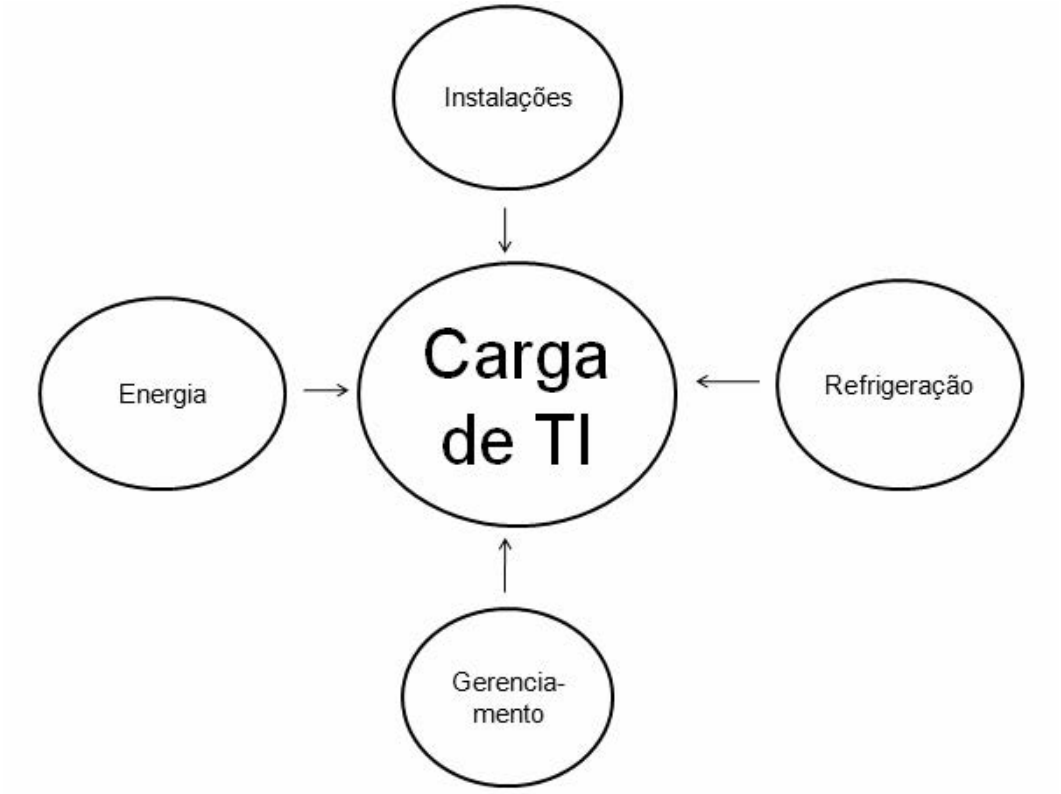


Figura 3-2 Componentes do DATACENTER

Com os usuários da organização solicitando cada vez mais capacidade de processamento e de armazenamento, os DATACENTERS que estão em operação atingiram a capacidade máxima, pois foram construídos há alguns anos. Os gerentes de TI estão avaliando novas tecnologias para utilizar melhor os recursos, incluindo também a busca por maior eficiência energética.

As organizações estão repensando a arquitetura e o planejamento do DATACENTER devido ao aumento destas demandas de processamento e armazenamento e também devido ao aumento das normas e da fiscalização rigorosa sobre a recuperação destas estruturas em caso de desastres. Estas novas demandas acabam provocando grande impacto nos níveis de serviço fornecidos e nos custos das operações baseadas no DATACENTER.

A consolidação de muitos DATACENTERS em poucos é uma resposta a estas novas demandas e se torna possível devido aos grandes avanços dos últimos dez anos em disponibilidade e confiabilidade de banda e de tecnologia de INTERNET.

A consolidação do DATACENTER envolve a migração de todos os ativos de TI, pessoas e as aplicações que recebem suporte, portanto um planejamento abrangente é crítico para que o negócio continue a funcionar sem interrupção.

O papel principal dos componentes do DATACENTER no novo cenário é possibilitar o atingimento do nível de serviço adequado para cada aplicação. A ideia é oferecer níveis de serviço de acordo com a criticidade da aplicação.

Os DATACENTERS hospedam recursos computacionais críticos em um ambiente necessariamente controlado e sob gerenciamento centralizado que permite suportar as aplicações empresariais.

É comum pensar o DATACENTER no nível dos serviços que o compõe como processamento, armazenamento, conectividade, segurança, gerenciamento, VIRTUALIZAÇÃO, alta disponibilidade e recuperação a desastres, automação e gerenciamento, conforme ilustra a **Figura 3-3**.

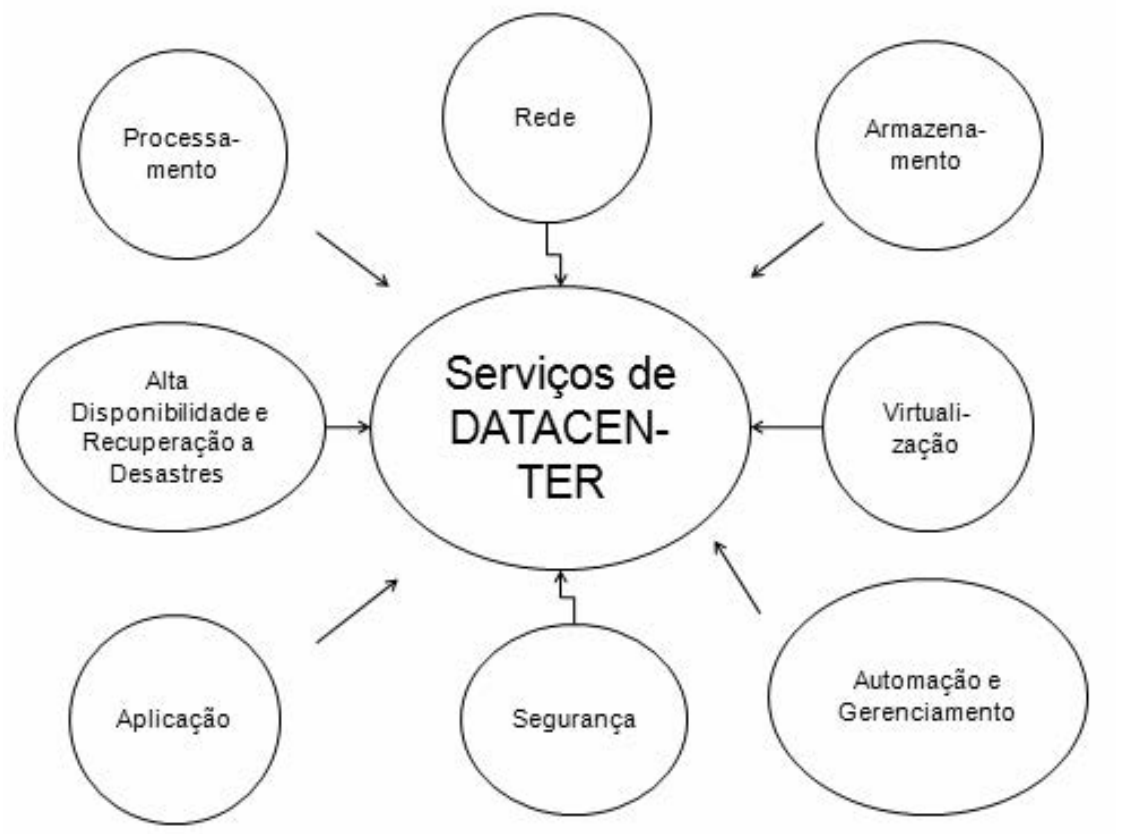


Figura 3-3 Serviços do DATACENTER

Os serviços de DATACENTER dependem dos componentes do DATACENTER. A meta central de um projeto de DATACENTER é obter resiliência, que, em TI, significa atender a demanda de negócios de maneira efetiva, reduzir o custo total de propriedade (TCO) e tornar o negócio flexivel. Os critérios a serem considerados em um projeto de DATACENTER para o atingimento desta meta são:

- Desempenho.
- Disponibilidade.
- Escalabilidade.
- Segurança.
- Gerenciabilidade.

Estes critérios devem ser utilizados no dimensionamento dos diversos dispositivos que compõem o DATACENTER e definem a qualidade dos serviços a serem providos para as aplicações. Estes serviços também devem ser projetados para funcionar em conjunto, ou seja, integrados, e atender as demandas das aplicações que, por sua vez, atendem as demandas dos processos de negócio.

Uma visão mais adequada do DATACENTER deve pensá-lo em termos de serviços. A qualidade dos dispositivos utilizados irá influenciar os níveis de serviço que serão entregues às aplicações e aos processos de negócio.

A **Figura 3-4** ilustra de maneira didática os serviços principais providos pelo DATACENTER, separando os serviços de TI dos componentes externos do DATACENTER como instalações, energia, refrigeração e gerenciamento físico para facilitar o entendimento. Alguns serviços acabam dependendo de outros e alguns serviços acabam completando outros.

A VIRTUALIZAÇÃO é o principal serviço de TI do DATACENTER e influenciará o dimensionamento dos demais serviços oferecidos. O correto projeto da estrutura de TI virtualizada é a chave do bom projeto de DATACENTER.

Justiificar um projeto de VIRTUALIZAÇÃO para uma infraestrutura nova ou mesmo já existente é uma tarefa fácil. As vantagens são inúmeras e o custo total de propriedade de uma solução virtualizada é facilmente demonstrado com a utilização de ferramentas disponibilizadas pelos fabricantes dos softwares de VIRTUALIZAÇÃO.

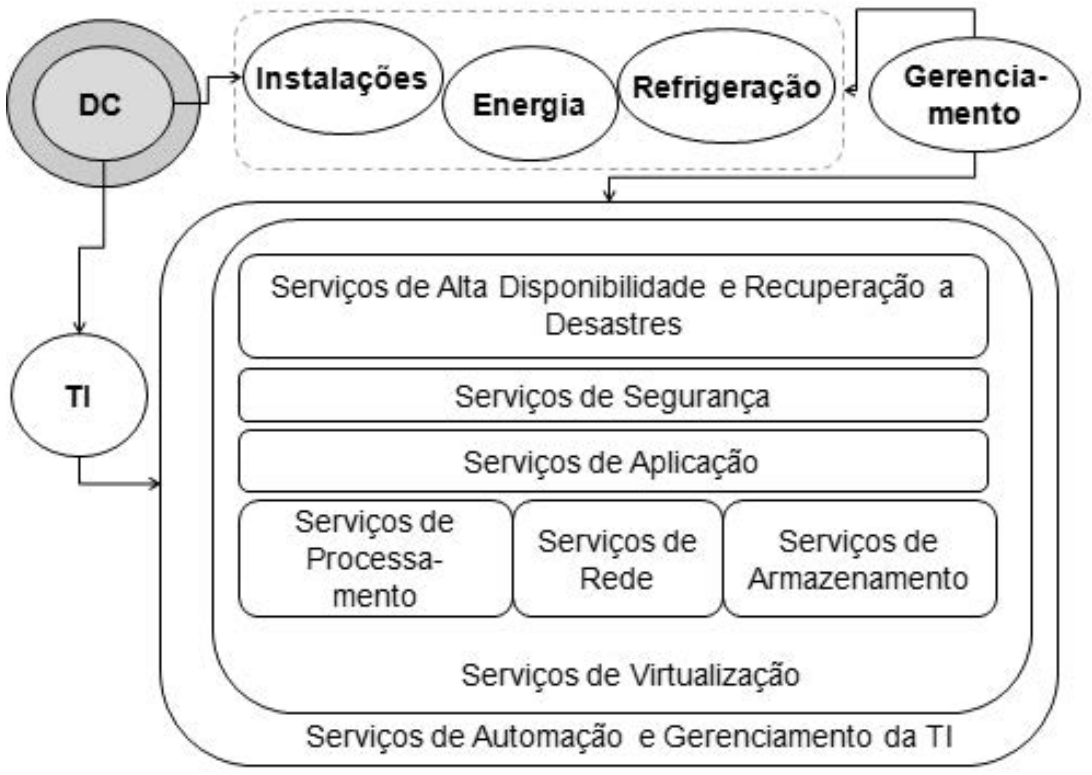


Figura 3-4 Serviços de DATACENTER

Os principais serviços do DATACENTER que permitem atingir os níveis de serviço necessários para as aplicações são:

- **Serviços de rede:** envolvem a conexão entre os componentes internos e deles com o mundo exterior. As conexões são feitas mediante a arbitragem dos switches. Os switches são dispositivos importantes de um projeto de conectividade IP e são utilizados em várias camadas para definir a hierarquia do DATACENTER.
- **Serviços de segurança:** serviços de segurança envolvem os serviços de firewall, que permitem o controle da navegação do usuário, negando acesso a aplicações e sites desnecessários, além de incorporar um sistema que permite aos gestores monitorar as atividades de cada usuário. Incluem também serviços providos por sistemas de detecção de intrusão e prevenção e filtragem de pacotes.
- **Serviços de processamento:** os serviços de processamento respondem diretamente pelo desempenho do DATACENTER. Os dispositivos envolvidos são os servidores, sistemas operacionais e processadores. Os servidores podem utilizar diversas arquiteturas, mas a tendência maior é a de utilizar a arquitetura x86.
- **Serviços de armazenamento:** envolvem o armazenamento de dados em unidades de storage. Redes de armazenamento e dispositivos de conexão ao storage são partes importantes da arquitetura. Também os níveis de serviço para disponibilidade e segurança são dependentes deste serviço.
- **Serviços de VIRTUALIZAÇÃO:** os serviços de VIRTUALIZAÇÃO permitem que servidores físicos rodem diversas aplicações em diferentes sistemas operacionais, otimizando a utilização dos recursos de processamento e memória. Sistemas operacionais de VIRTUALIZAÇÃO e suas funcionalidades são aspectos importantes de serem tratados.
- **Serviços de aplicação:** os serviços de aplicação envolvem *load-balancing*, *secure socket layer* (SSL), *offloading* e *caching*.
- **Serviços de alta disponibilidade (*high availability* – HA) e recuperação a desastres (*disaster recovery* – DR):** Tratam dos serviços para obtenção da alta disponibilidade e recuperação a desastres, incluindo a extensão da SAN, seleção do site de contingência, interconectividade. Tratam também de políticas, softwares e dispositivos de backup restore e da replicação.
- **Serviços de automação e gerenciamento:** envolvem toda a malha de gerenciamento incluindo o NOC (*networking operation center*), desde o hardware até os aspectos de automação de patches (correções) de sistema operacional. O gerenciamento deve possibilitar uma operação assistida 24x7 e executada até remotamente. O planejamento e o controle da produção devem garantir a execução dos processos de TI como as paradas programadas, execução de jobs e rotinas periódicas.

A **Figura 3-5** ilustra a relação entre os serviços de DATACENTER e os serviços de Infraestrutura de TI normalmente baseados na biblioteca ITIL. O nível de serviço percebido pelo usuário do processo depende do nível de serviço da aplicação, que, por sua vez, depende do nível de serviço da infraestrutura, onde os serviços de DATACENTER estão incluídos.

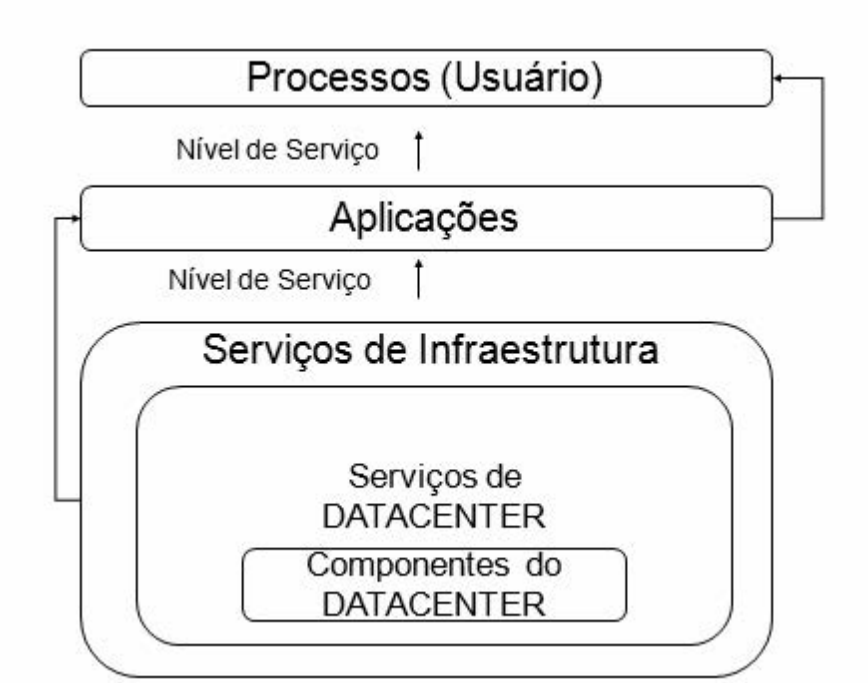


Figura 3-5 Serviços de infraestrutura de TI e serviços de DATACENTER

Os serviços de DATACENTER são interdependentes e devem ser implementados de forma integrada para possibilitar o atingimento do SLA entregue à aplicação. Ou seja, o projeto de um DATACENTER deve considerar todos estes serviços de forma a obter os níveis de serviço necessários às aplicações e consequentemente ao negócio.

3.2. TIER DATACENTER

A ideia de enxergar a montagem do DATACENTER como uma montagem de Lego induz a necessidade de construí-lo em pequenas partes integradas ou blocos padronizados. A norma TIA-942 (*Telecomunications Infrastructure for DATACENTERS*), de 2005, define um padrão de classificação para aspectos básicos do DATACENTER como:

- Layout e espaço físico;
- Infraestrutura de cabeamento;
- Condições ambientais;
- Classificação por níveis de disponibilidade (baseado na definição do *Uptime Institute*).

O desempenho e a disponibilidade do DATACENTER podem ser descritos por uma série de termos como confiabilidade, disponibilidade e MTBF, normalmente difíceis de serem calculados. Uma alternativa é categorizar o DATACENTER em termos de camadas (TIERS) de níveis críticos. Existe sempre um compromisso entre o custo total de propriedade (TCO) e a escolha de uma camada de criticidade. Ou seja, quanto mais sofisticado for o DATACENTER, menor o *downtime* e maior é o seu custo total de propriedade.

No projeto do DATACENTER, a etapa de planejamento requer muito cuidado, pois é onde normalmente se encontram os maiores erros. Deve-se construir uma especificação resultante das preferências e limitações do projeto.

A APC no relatório técnico 122 – *Guidelines for Specifying DATACENTER Critically/Tier Levels* – aponta os métodos de classificação de DATACENTER mais conhecidos, que são o *The Uptime Institute's Tier Performance Standards*, TIA 942 *e Syska Henessy Group's Criticality Levels*.

- ***Uptime Performance Standard:*** criado em 1995, é muito utilizado em projetos de DATACENTERS, mas não especifica detalhes de projeto. A criticidade deve ser sempre obedecida em todos os critérios específicados.
- **ANSI/EIA/TIA 942:** baseado nas quatro camadas do *Uptime Performance Standard*, já está na revisão 5, de 2005. Será detalhada um pouco mais adiante por ser a mais utilizada.
- ***Syska Henessy Group:*** baseado em dez camadas que se relacionam com as quatro camadas do *Uptime Performance Standard*.

A norma ANSI/EIA/TIA 942 (ou TIA 942) permite definir o nível de criticidade do DATACENTER e é, sem dúvida, o padrão utilizado para classificar DATACENTERS no mundo. A utilização da norma deve considerar se está se tratando de um novo projeto (*greenfield* nos EUA) ou um DATACENTER já existente (*retrofit* nos EUA). Para DATACENTERS de provedores é normal que o nível de criticidade seja elevado, pois os custos envolvidos para a obtenção de um alto nível de criticidade são diluídos. No caso de DATACENTERS próprios, esta é uma questão complexa, pois normalmente os custos para obtenção de um alto nível de criticidade são muito altos.

A norma TIA 942 indica os requisitos desde a construção até a pronta ativação do DATACENTER. Esta norma é baseada em um conjunto de outras normas relacionadas abaixo:

- **ASHRAE:** trata da refrigeração.
- **TIA/EIA 568:** estabelece um sistema de cabeamento para aplicações genéricas de telecomunicações em edifícios comerciais. Permite o planejamento e a instalação de um sistema de cabeamento estruturado.
- **TIA/EIA 569:** norma que trata dos encaminhamentos e espaços.
- **TIA/EIA 606:** providencia um esquema de administração uniforme independente das aplicações.
- **TIA/EIA 607:** trata das especificações de aterramento e links dos sistemas de cabeamento estruturado em prédios comerciais. Providencia especificações sobre aterramento e links relacionados à infraestrutura de telecomunicações do edifício.

A topologia básica de um DATACENTER pode ser descrita de várias formas. A descrição baseada na norma TIA 942 é uma espécie de padrão para este tipo de ambiente. A classificação em camadas TIA 942 é muito utilizada para comparação de DATACENTERS, estimativa de custo e desempenho.

IMPORTANTE: O *rating* da camada só é mantido se os aspectos chaves da camada forem considerados. Ou seja, não adianta ter um DATACENTER robusto em termos elétricos e um sistema inadequado de refrigeração, pois, para efeito de classificação, o que vale é o sistema inadequado de refrigeração.

Principais componentes:

- ***Entrance Room* (ER):** Sala de Entrada. Espaço de interconexão entre o cabeamento estruturado do DATACENTER e o cabeamento vindo das operadoras de telecomunicações.
- ***Main Distribution Area* (MDA):** Área onde se encontra a conexão central do DATACENTER e de onde se distribui o cabeamento estruturado. Inclui roteadores e o backbone.
- ***Horizontal Distribution Area* (HDA):** Aréa utilizada para conexão com área de equipamentos. Inclui o cross conect horizontal e equipamentos intermediários. Incluem switches LAN/SAN/KVM.
- ***Zone Distribution Area* (ZDA):** Ponto de interconexão opcional do cabeamento horizontal. Provê flexibilidade para o DATACENTER. Fica entre o HDA e o EDA.
- ***Equipment Distribution Area* (EDA):** Área para equipamentos terminais (servidores, *storage*, unidades de fita) e os equipamentos de rede. Inclui RACKS e gabinete.

A **Figura 3-6** ilustra os componentes do modelo de DATACENTER em camadas:

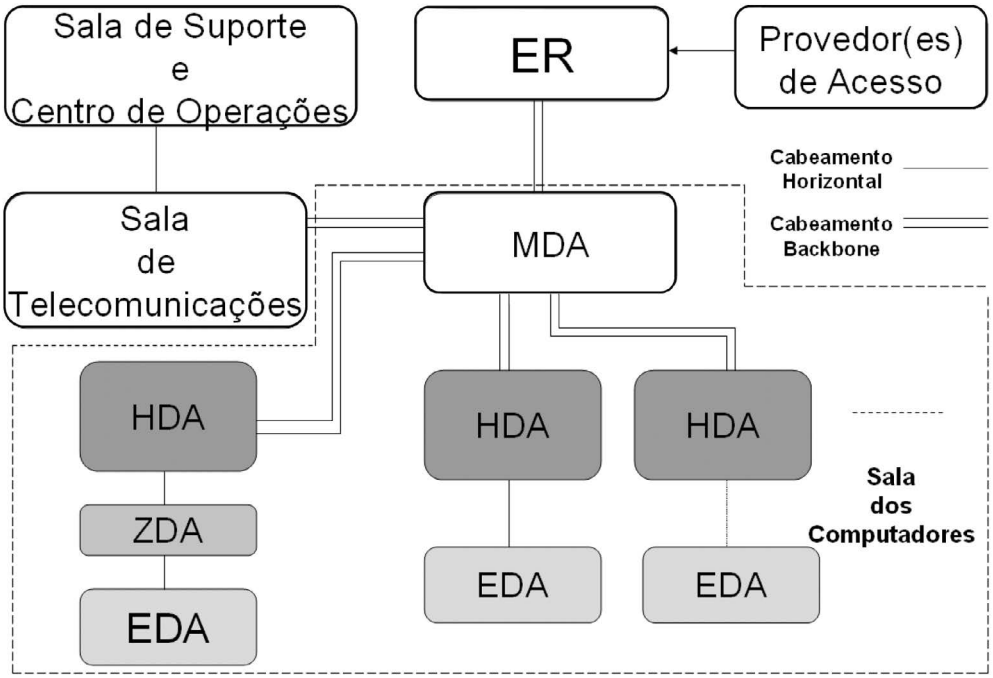


Figura 3-6 Topologia básica de um DATACENTER segundo a TIA 942

Os servidores físicos são colocados em RACKS e a sua interligação feita por switches de acesso que, por sua vez, se interligam a switches de agregação através de uplinks de 10Gbps como padrão. Os switches de acesso em geral são do tipo ToR (Top of Rack) e os switches de agregação são do tipo EoR (End of Row).

A norma TIA942[4] define a classificação dos DATACENTERS em quatro níveis independentes, chamados de TIERS (camadas), considerando Arquitetura, Telecomunicações, Aspectos Elétricos e Mecânicos.

3.2.1. TIER 1

Básico: neste modelo não existe redundância nas rotas físicas e lógicas. Prevê a distribuição de energia elétrica para atender a carga sem redundância. A falha elétrica pode causar interrupção parcial ou total das operações.

- **Pontos de Falha**: falta de energia da concessionária no DATACENTER ou na Central de Operadora de Telecomunicações. Falha de equipamentos da Operadora. Falha nos roteadores ou switches quando não redundantes. Qualquer catástrofe nos caminhos de interligação ou nas áreas ER, MDA, HDA, ZDA e EDA.
- ***Downtime:*** 28.8 hr/ano (equivale a 99.671%).

3.2.2. TIER 2

Componentes Redundantes: os equipamentos de telecomunicações do DATACENTER e também os equipamentos de telecomunicações da operadora, assim como os dispositivos da LAN-SAN, devem ter módulos de energia redundantes. O cabeamento do backbone principal LAN e SAN das áreas de distribuição horizontal para os switches do backbone deve ter cabos de cobre ou fibras redundantes. Na TIER 2 deve-se ter duas caixas de acesso de telecomunicações e dois caminhos de entrada até o ER. Deve-se prover módulos de NO-BREAK redundantes em N+1. É necessário um sistema de gerador elétrico para suprir toda a carga. Não é necessário redundância na entrada do serviço de distribuição de energia. Os sistemas de ar condicionado devem ser projetados para operação contínua 24x7 e incorporam redundância N+1.

- **Pontos de Falha**: falhas nos sistemas de ar condicionado ou de energia podem ocasionar falhas em todos os outros componentes do DATACENTER.
- ***Downtime:*** 22 hr/ano (equivale a 99.749%).

3.2.3. TIER 3

Sistema Autossustentado: deve ser atendido por, no mínimo, duas operadoras de telecomunicações com cabeamentos distintos. Deve-se ter duas salas de entrada (ER) com, no mínimo, 20 metros de separação. Estas salas não podem compartilhar equipamentos de telecomunicações e devem estar em zonas de proteção contra incêndios, sistemas de energia e ar condicionado distintos. Deve-se prover caminhos redundantes

4 Ver texto DATACENTER em http://www.furukawa.com.br/

entre as salas de entrada (ER), as salas MDA e as salas HDA. Nestas conexões, deve-se ter fibras ou pares de fios redundantes. Deve-se ter uma solução de redundância para os elementos ativos considerados críticos como o *storage*. Deve-se prover pelo menos uma redundância elétrica N+1.

- **Pontos de Falha:** qualquer catástrofe no MDA ou HDA irá interromper os serviços.
- ***Downtime***: 1.6 hr/ano (equivale a 99.982%).

3.2.4. TIER 4

Sem Tolerância a Falhas: todo o cabeamento deve ser redundante, além de protegido por caminhos fechados. Os dispositivos ativos devem ser redundantes e devem ter alimentação de energia redundante. O sistema deve prover comutação automática para os dispositivos de backup. Recomenda-se uma MDA secundária que, em zonas de proteção contra incêndio, devem ser separadas e alimentadas por caminho separado. Não é necessário um caminho duplo até o EDA. Deve-se prover disponibilidade elétrica 2(N+1). O prédio deve ter pelo menos duas alimentações de energia de empresas públicas, a partir de diferentes subestações. O sistema de ar condicionado deve incluir múltiplas unidades de ar condicionado com a capacidade de resfriamento combinada para manter a temperatura e a umidade relativa de áreas críticas nas condições projetadas.

- **Pontos de falha**: se a MDA primária falhar e não houver MDA secundária. Se a HDA primária falhar e não houver HDA secundária.
- ***Downtime:*** 0.4 hr/ano (equivale a 99.995%).

A **Tabela 3-1** resume o *downtime* sugerido para cada camada da norma TIA 942 com um indicativo de custo ($/W) de construção nos Estados Unidos. Sair de um modelo TIER 1 para TIER 3 em um DATACENTER DE 1000KW significa sair de um projeto de US$ 10M para um projeto de US$ 20M. O DATACENTER padrão dos grandes provedores é o TIER 2, mesmo fora do Brasil. Também não adianta o DATACENTER ter um elevado nível de disponibilidade (TIER 3/4) se os equipamentos de TI apresentam baixo nível de disponibilidade. Recentemente o DATACENTER da ATIVAS, no Brasil, recebeu a certificação UPTIME de TIER 3, sendo o primeiro da América Latina.

Tabela 3-1 TIA 942 – *Downtime*, *Uptime* e Custo por Camada

	Ano de Surgimento	*Downtime* (hr/ano)	Uptime (%)	Custo ($/W)
TIER 1	1965	28.8	99.671	10
TIER 2	1970	22	99.749	11
TIER 3	1985	1.6	99.982	20
TIER 4	1995	0.4	99.995	22

A HP recentemente propôs o modelo de DATACENTER híbrido, um DATACENTER construído com diferentes níveis de criticidade de forma a otimizar a relação entre custo e *downtime*. A **Figura 3-7** ilustra a ideia central.

Figura 3-7 DATACENTER Híbrido

3.3. TCO do DATACENTER

Não existe uma metodologia padrão para cálculo do custo total de propriedade (TCO) do DATACENTER. De uma forma geral, pode-se afirmar que o TCO de um DATACENTER é dado pela seguinte fórmula:

TCO (DC) = DATACENTER depreciação + DATACENTER Opex + TI depreciação + TI Opex.

Na fórmula os custos do DATACENTER relacionados à construção das instalações, equipamentos de energia e refrigeração (DATACENTER depreciação e DATACENTER Opex) estão separados dos custos da carga de TI (TI depreciação e TI Opex) que envolvem servidores, storage, redes e softwares.

O custo de capital (Capex) do DATACENTER varia largamente e depende da localização, projeto, velocidade da construção, classificação em TIER, etc. Grandes DATACENTERS custam em média $12-$15/W para construir nos EUA. A depreciação utilizada em DATACENTER é tipicamente de 10 -15 anos.

Em instalações de DATACENTER, 80% em média é custo com os equipamentos de energia e refrigeração e 20% é custo com a construção em si. Em instalações convencionais de edifícios a relação é o contrário.

Os objetivos para energia e densidade de calor (total KW, KW/cabinet ou W/m^2) e o nível de disponibilidade (TIER) afetam diretamente o custo de capital do DC junto com o tamanho do espaço físico.

O custo operacional (Opex) também é muito dificil de ser calculado, pois depende dos padrões utilizados. Nos Estados Unidos, este custo sem o custo da energia é da or-

dem de $0,02W-$0,08W por mês. O custo operacional dos equipamentos de TI também varia muito e depende dos contratos acordados.

Estudos relacionados pela APC sugerem custos de Opex e Capex equivalentes para DATACENTERS em um típico ciclo de vida. Segundo a APC, a maior fonte de melhoria de TCO do DATACENTER é fazer o correto dimensionamento dos equipamentos de energia e refrigeração. A energia pode representar de 20% a 40% do TCO de um DATACENTER.

3.4. Padronização do DATACENTER

Os principais *form-factors* utilizados em DATACENTERS são o CHASSI, o RACK e o CONTAINER, descritos a seguir.

3.4.1. CHASSI

Possibilita construir os DATACENTERS baseados em blocos do tipo CHASSIS, que utilizam servidores em lâminas (blades), conforme ilustra a **Figura 3-8**.

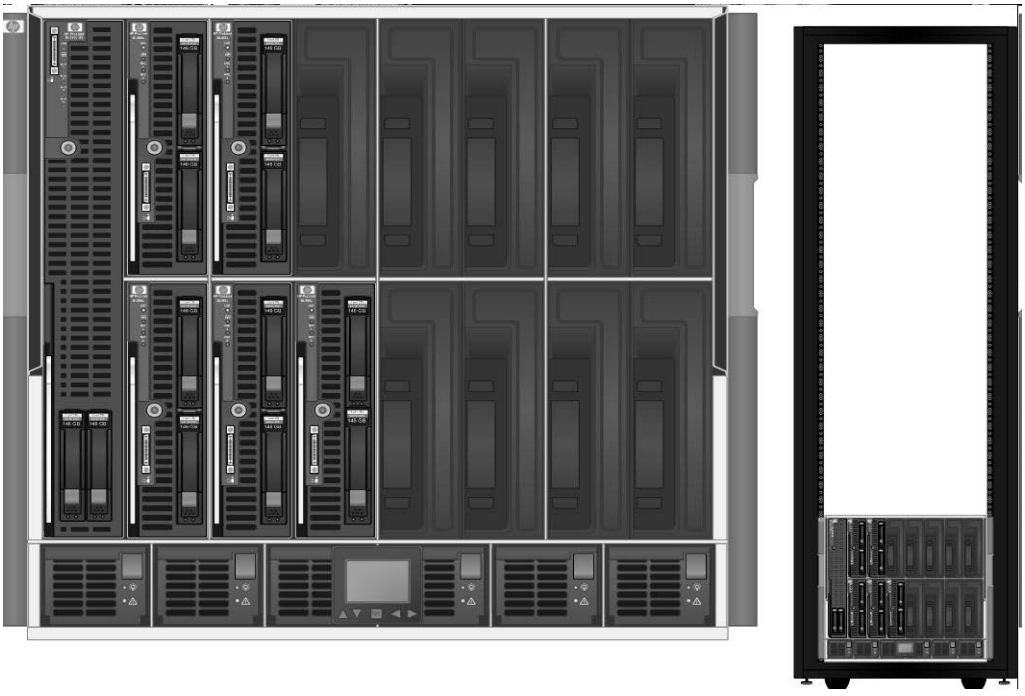


Figura 3-8 Chassi Blade c7000 da HP

3.4.2. RACK

Possibilita construir os DATACENTERS da forma convencional em blocos de RACKS de servidores, conforme ilustra a **Figura 3-9**, permitindo obter uma modularidade média sem o compromisso de utilizar chassis de blades para o *form-factor* de todos os servidores. Os RACKS são construídos em três principais tamanhos: 24U, 42U e 48U, onde 1U = 44,45mm.

Figura 3-9 RACKS padrões 24U e 42U da Dell

A **Figura 3-10** ilustra as dimensões do RACK Dell

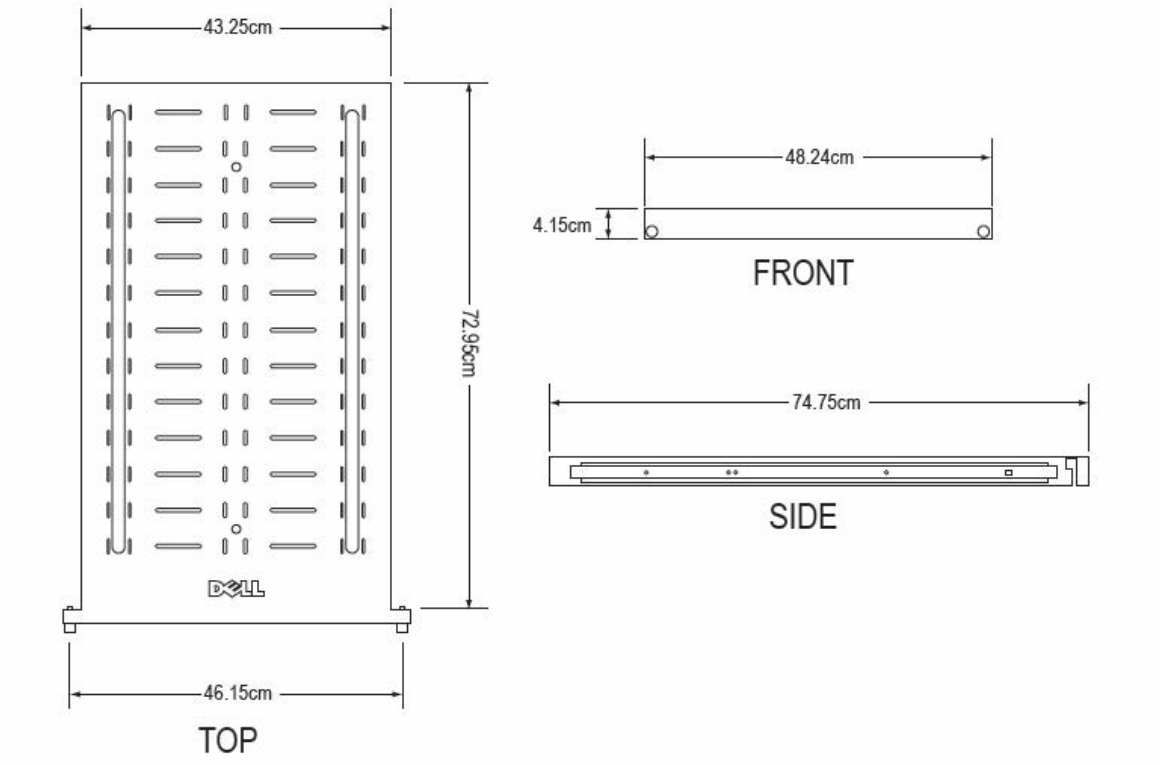


Figura 3-10 Dimensões do RACK DELL

3.4.3. CONTAINER

Os RACKS de servidores podem ser embutidos em CONTAINERS que permitem o fácil deslocamento entre prédios, facilitando a montagem e permitindo a escalabilidade de equipamentos de energia, processamento e armazenagem em uma dimensão mais ampla.

Empresas como a Activepower estão portando os equipamentos de energia (no-breaks e geradores) e de resfriamento de água (*chillers*) em CONTAINERS específicos. A solução PowerHouse, por exemplo, fornecida pela Activepower, é integrada com a solução CONTAINER da Sun, o MD-20, que permite ter RACKS internos com até 25KW de potência. Neste caso, seriam necessários dois CONTAINERS, um para TI e outro para energia e refrigeração.

A solução PowerHouse utiliza no-breaks baseados na tecnologia ***flywheel*** que dispensam o uso de baterias e podem alimentar CONTAINERS de até 1MW. Os DATACENTERS baseados em CONTAINERS normalmente utilizam entre 1000 e 2000 servidores formando um POD (*performance optimized datacenter*) que só necessita de energia, refrigeração e banda.

Os CONTAINERS são interessantes para DATACENTERS que podem ser disponibilizados rapidamente em eventos, para sites de contingência, em megainstalações que precisam ser escaláveis e até como DATACENTERS principais. Empresas como HP, IBM, Sun e Cisco já fornecem estas estruturas, e Microsoft e Google atualmente utilizam CONTAINERS próprios.

A **Figura 3-11** ilustra o CONTAINER S20/D20 da Sun. A principal vantagem destas estruturas é possibilitar o aumento da eficiência energética, basicamente a relação entre a energia que é adquirida da concessionária e a energia que efetivamente é consumida pela carga de TI.

A diferença na forma de fazer a capacidade de processamento de um DATACENTER baseado em CONTAINER crescer, quando comparado a uma solução convencional, é que o CONTAINER permite um crescimento mais granular em uma escala de tempo menor.

A **Figura 3-12** ilustra a vantagem do CONTAINER e mostra que o crescimento em um arranjo baseado em CONTAINERS pode acontecer em intervalos menores.

Os CONTAINERS permitem rápida expansão e flexibilidade quando comparados aos DATACENTERS convencionais. Também a sua modularidade permite reduzir riscos no planejamento do crescimento futuro.

Figura 3-11 CONTAINER Sun S20/D20

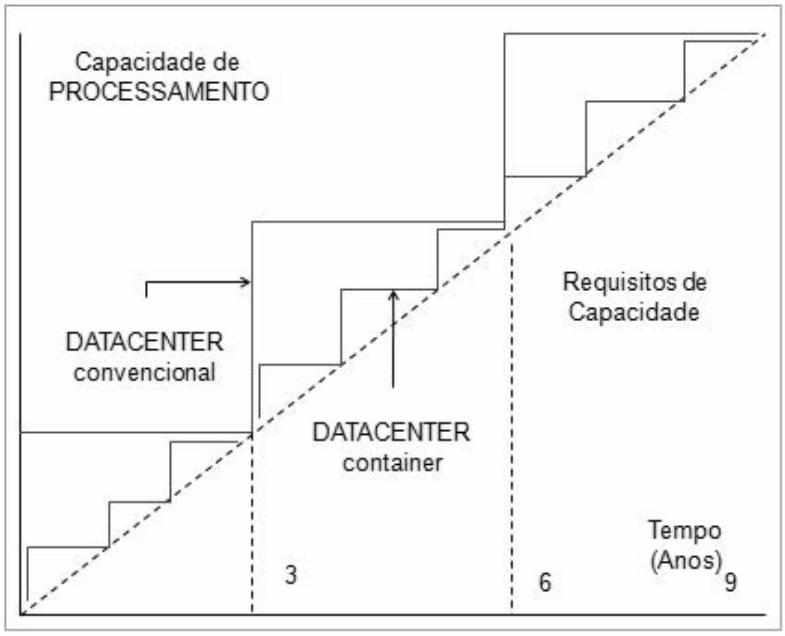


Figura 3-12 Efeito CONTAINER na Escalabilidade

3.4.4. Conceito na Prática: DATACENTER POD da HP

A família HP POD de CONTAINERS é disponibilizada em dois tamanhos: 6 metros e 12 metros. O CONTAINER HP 2000C é o de seis metros e vem com 500U de capacidade de RACK para dez RACKS de 50U, possui capacidade de 290KW sem redundância

(29 KW por RACK) e 145KW (14.5 KW por RACK) com redundância N+N. Podem ser instalados 1600 nós usando os blades BL2x220C. Ou seja, são 80 blades por RACK que totalizam 800 blades e 1600 nós em uma configuração máxima.

O HP 2000C suporta hardware padrão de terceiros e utiliza RACKS convencionais de 482 mm de profundidade. O fabricante aponta que o PUE do HP 2000C, parâmetro a ser explicado ainda neste capítulo, é da ordem de 1.25, mesmo considerando o consumo do *chiller*. O projeto de refrigeração é do tipo "closely coupled".

A **Figura 3-13** ilustra o container HP 2000C.

Figura 3-13 CONTAINER HP 2000C

3.4.5. Conceito na Prática: SALA COFRE da Aceco

Em construções convencionais de DATACENTERS a SALA COFRE permite obter um ambiente estanque que protege o DATACENTER contra ameaças físicas incluindo fogo, calor, umidade, água, fumaça, arma de fogo e acesso indevido. A TI é o que fica dentro da SALA COFRE, não é a SALA COFRE. O papel da SALA COFRE é proteger as informações e os equipamentos de TI.

Normalmente a SALA COFRE é construída de forma escalável, podendo ser ampliada ou mudada para outro local. A **Figura 3-14** ilustra uma típica SALA COFRE. A SALA COFRE deve ser testada integralmente de acordo com normas específicas de segurança.

O procedimento de certificação da ABNT/INMETRO PE047.01 regulamenta e exige a execução de diversos testes, segundo normas específicas que garantem o desempenho e a conformidade da SALA COFRE com as normas NBR 11515 e ISSO 27002. O Procedimento PE047.01 inclui também uma série de testes de fogo, calor e umidade para mitigar outros riscos existentes baseado na norma NBR 15427. Importante ressaltar que a SALA COFRE não é solução para projetos baseados em CONTAINERS.

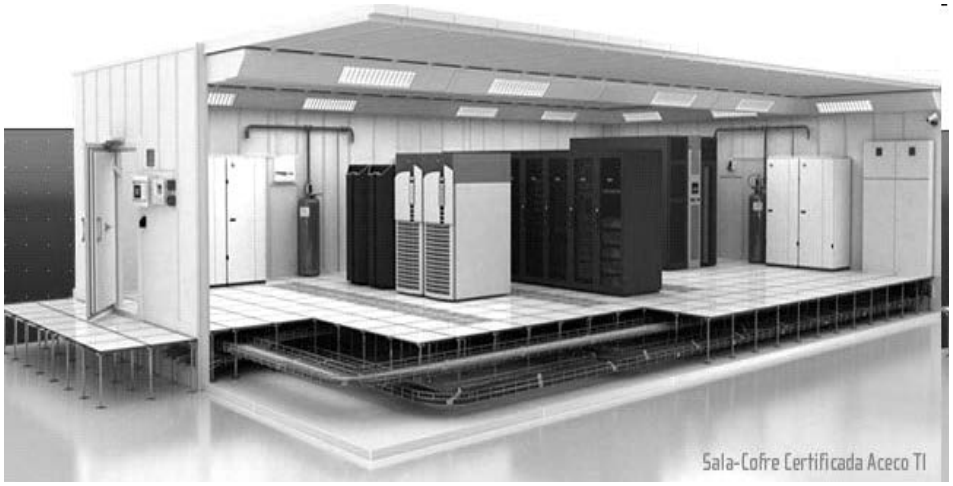


Figura 3-14 SALA COFRE da Aceco TI

3.5. Construção do DATACENTER

A construção do DATACENTER envolve disponibilização de energia e refrigeração, cabeamento, controles de umidade e temperatura, sistemas de prevenção contra incêndios, segurança física, espaço para os *RACKS* e piso elevado.

- **Energia e Refrigeração**: o provimento de energia e refrigeração deve estar de acordo com as premissas adotadas quando da formulação do projeto. A carga de TI e a carga dos equipamentos que suportam a refrigeração e o seu crescimento futuro devem ser consideradas no correto dimensionamento dos componentes chaves do DATACENTER.
- **Temperatura e Umidade**: devem ser monitoradas e controladas por meio de sistemas de climatização precisos e redundantes. Os DATACENTERS utilizam em muitos casos pisos elevados para possibilitar que o fluxo de ar seja direcionado para baixo dos equipamentos, a fim de mantê-los refrigerados.
- **Segurança Física**: o acesso do pessoal no DATACENTER deve ser controlado. Minimizar a quantidade de pessoas que podem entrar no DATACENTER e melhorar o controle é uma medida para mitigar risco. A segurança física pode ser pensada em termos de profundidade de segurança (*depth of security*), ou seja, é necessário pensar a segurança em diferentes níveis, partindo de níveis mais leves de segurança, no estacionamento, por exemplo, para níveis mais fortes de segurança próximos à carga de TI. A Figura 3-15 ilustra a segurança de perímetro no DATACENTER.

Figura 3-15 Segurança Física do DATACENTER

- **Gerenciamento e Monitoração:** o gerenciamento e a monitoração de todo o ambiente também são aspectos relevantes. O Network Operation Center (NOC) é a célula central do gerenciamento. Por definição, o NOC é um local onde se centraliza a gerência da rede e dos componentes do DATACENTER. A partir desse centro e de aplicações que monitoram a rede os administradores podem saber, em tempo real, a situação de cada "componente" dentro da rede. Os componentes são os servidores, roteadores, switches, storage, etc. O NOC normalmente é utilizado em grandes DATACENTERS , mas a tendência é que as pequenas e médias empresas também adotem este tipo de gerenciamento. O NOC pode ser inclusive terceirizado com um provedor que dilua o seu custo com a venda do serviço de NOC para pequenas e médias empresas.

3.6. Impactos da VIRTUALIZAÇÃO no DATACENTER

A VIRTUALIZAÇÃO é a tecnologia que permitiu que o DATACENTER pudesse alocar recursos em função da demanda. A otimização propiciada pela VIRTUALIZAÇÃO trouxe inúmeras vantagens, incluindo a redução do espaço físico, menor consumo de energia, simplificação do ambiente de recuperação a desastres e melhora do nível de disponibilidade das aplicações.

O ambiente virtualizado fornece maior flexibilidade operacional à infraestrutura e dinamiza as alterações nos sistemas, enquanto fornece uma plataforma que propicia mais facilmente obter continuidade do negócio e escala rapidamente para satisfazer as demandas da empresa. Ambientes virtualizados melhoram a capacidade da organização de TI de fornecer soluções potentes e robustas para vencer os desafios de forma rápida e eficiente. O departamento de TI não precisa mais configurar cada elemento do sistema para que todos funcionem juntos, o que reduz a complexidade da infraestrutura de TI.

O gerenciamento da infraestrutura física e virtual com um conjunto de ferramentas unificado elimina as dificuldades da integração de sistemas distintos. Os recursos de TI podem concentrar-se na adição de recursos avançados e na personalização das soluções para atender melhor às necessidades da empresa.

3.7. Referências Bibliográficas

Aceco TI. Catálogo de Soluções, 2009.

Cost Model for Planning, Development and Operation of a DATACENTER. Chandrakant D. Patel, Amip J. Shah, Internet Systems and Storage Laboratory, HP Laboratories Palo Alto. HPL-2005-107(R.1), June 9, 2005.

Guidelines for Specifying DATACENTER Critically/Tier Levels. APC White Paper #122.

Hornby, David & Pepple, Ken. Consolidation in the DATACENTER: Simplifying IT Environments to Reduce Total Cost of Ownership. Prentice Hall-PTR, 2002.

IT Consolidation: Business Drivers, Benefits and Vendor Selection, IDC Executive Brief, 2005.

Jayaswal, Kailash. Administering DATACENTERS: Servers, *Storage*, and Voice over IP. Wiley, 2005.

Luiz André, Barroso & Hölzle, Urs. The DATACENTER as a Computer; an Introduction to the Design of Warehouse-Scale Machines, 2009.

Opções atuais para consolidação, EMC, 2005.

Physical Security in Mission Critial Facilities. APC White Paper #82.

Saad, Alfredo C. Terceirização de Serviços de TI. Brasport, 2006.

Self-Contained Computer Room in a Shipping CONTAINER from Sun Microsystems. Uptime Institute, 2008.

Tier Classification define Site Infrastructure Performance, Uptime Institute, 2005.

Uma arquitetura aperfeiçoada para DATACENTERS de alta densidade e alta eficiência. APC White Paper #126.

Veras, Manoel. DATACENTER: Componente Central da Infraestrutura de TI, Brasport, 2009.

Verdi, Fábio Luciano; Rothenberg, Christian & Pasquini, Rafael, Magalhães, Maurício Ferreira. Novas Arquiteturas de DATACENTER para Cloud Computing, 2010.

4. Green Datacenter e Virtualização

4.1. Introdução

Para entender o que é o movimento verde (*green*), precisa-se antes saber o que é sustentabilidade empresarial. Sustentabilidade empresarial é um conceito sistêmico relacionado à continuidade dos aspectos econômicos, sociais, culturais e ambientais da sociedade humana. Precisa-se considerar que os aspectos socioambientais exercem influência crescente nas decisões das empresas e que o desenvolvimento sustentável é a ideia de não esgotar os recursos para o futuro.

O aspecto ambiental é o que está diretamente relacionado com o verde. É necessário considerar aspectos ambientais no projeto de novas estruturas e na operação da infraestrutura de TI, considerando o grande consumo de energia provocado por estas estruturas que são cada vez mais densas e, portanto, grandes geradoras de calor. O DATACENTER é um elemento chave da infraestrutura e onde acontece boa parte do consumo de energia do ambiente de TI. Projetá-lo obedecendo a aspectos de maior eficiência energética é o aspecto relevante do momento.

4.2. Eficiência do DATACENTER

A eficiência do DATACENTER, há bem pouco tempo, era medida unicamente em termos de indicadores vinculados à disponibilidade e ao desempenho. Com os aspectos ambientais sendo cada vez mais considerados, o aumento dos custos de energia e a limitação no fornecimento de energia por parte de alguns provedores de energia, é natural que os gerentes de infraestrutura de TI repensem a estratégia para o DATACENTER e considerem o aspecto do verde nas diversas escolhas que precisam fazer incluindo engenharia, equipamentos, tecnologias e a própria operação.

Estudos realizados na Universidade de Stanford indicam que o consumo de energia dos DATACENTERS representa 1,2% de todo o consumo de eletricidade nos EUA. O *Uptime Institute* levantou que os custos de energia representam, hoje, até 44% do TCO de um DATACENTER. Com o aumento dos custos com energia, provocado por servidores cada vez mais densos, é natural que a responsabilidade pelos gastos com energia do DATACENTER deve ir para dentro do departamento de TI e, neste momento, sugiro que você já tenha adequado a sua estratégia verde ao DATACENTER.

A eficiência do DATACENTER em operação, construído há mais de quinze anos, em geral não passa de 40%. Ou seja, de 100% de energia que é injetada no DATACENTER, só 40% alimentam a carga de TI. Em média, 60% da energia que alimenta o DATACENTER, na verdade, é consumida antes de chegar à carga de TI.

A **Figura 4-1** ilustra as fontes de consumo de energia no DATACENTER típico baseado em referência da APC. A indústria está fazendo um esforço enorme para mudar este quadro, e novos projetos de DATACENTER consideram novas formas de fornecimento de energia e otimização do tamanho dos equipamentos de energia e refrigeração. A forma de resfriar o DATACENTER é, sem dúvida, o grande gargalo que provoca boa parte da ineficiência do DATACENTER.

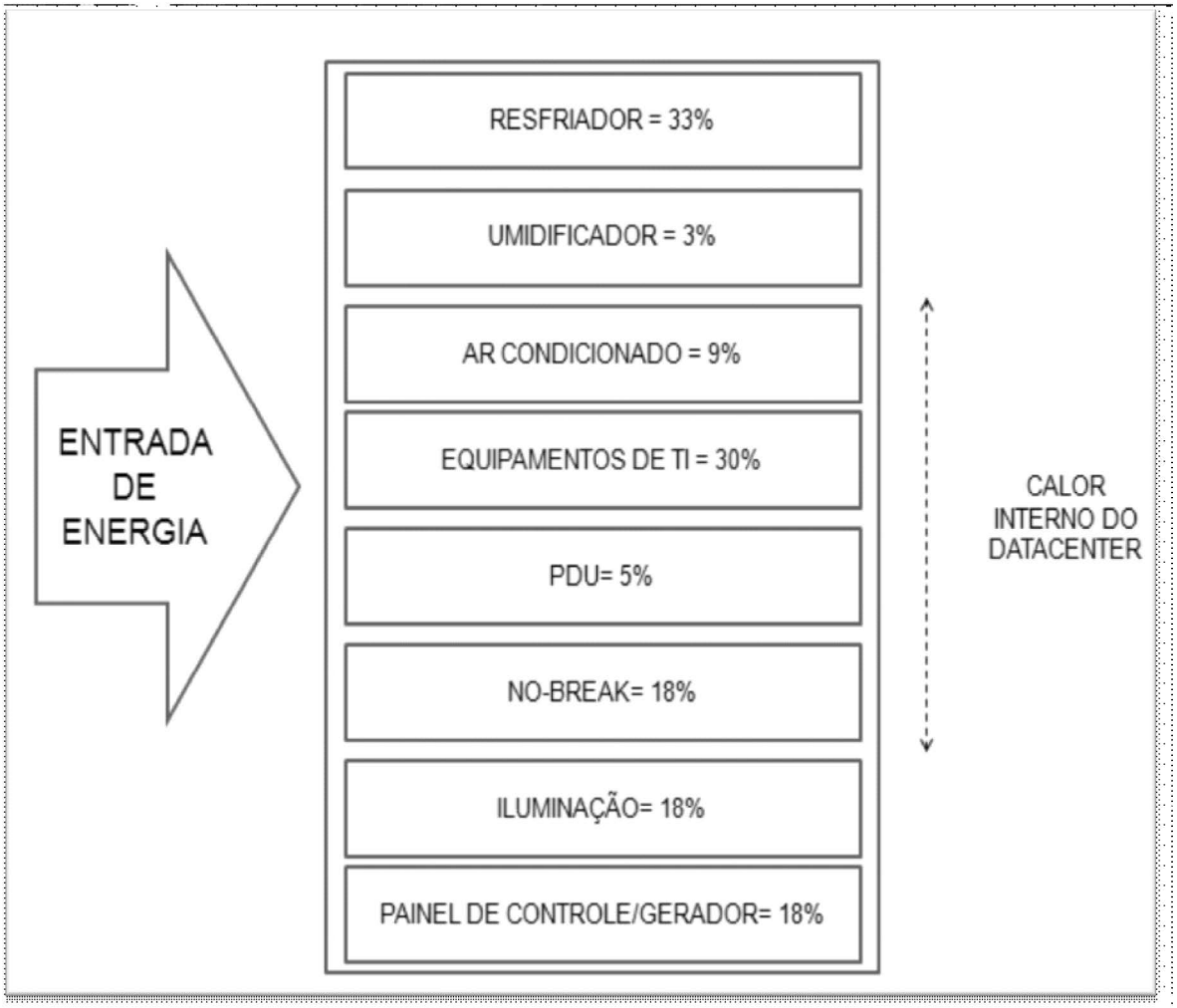


Figura 4-1 Fontes de Consumo de Energia em um DATACENTER típico

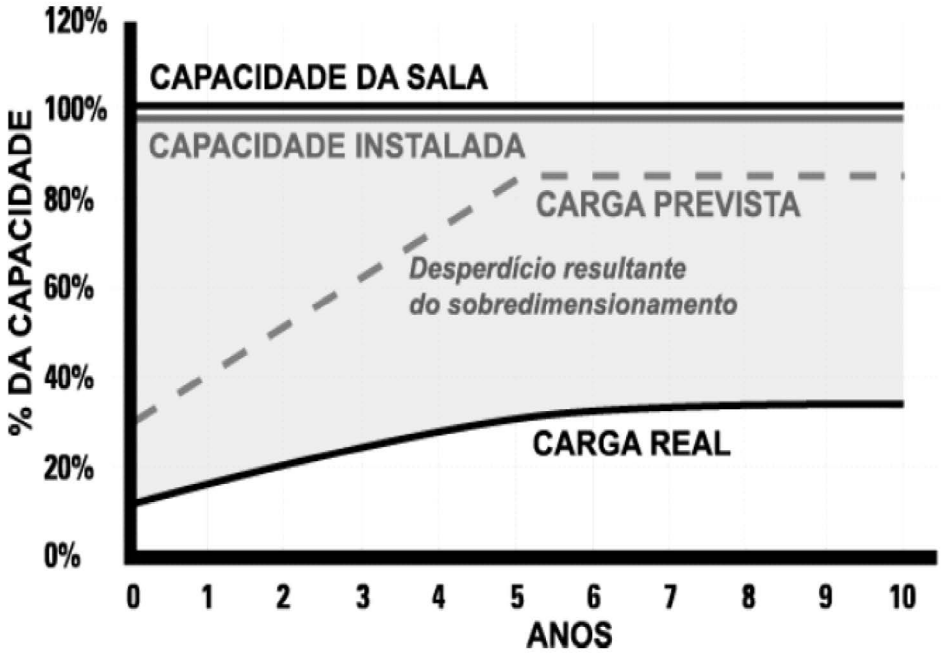
O artigo *DATACENTER Efficiency in the Scalable Enterprise*, publicado pela Dell Power Solutions, em 2007, mostra que a Dell monitorou em laboratório o consumo de um DATACENTER típico, construído há alguns anos, e verificou que de 100% da energia injetada no DATACENTER, só 41% efetivamente foram utilizados pela carga de TI, 28% foram consumidos pelos equipamentos que transformam e adequam a energia que chega

à carga de TI e 31% foram consumidos pelos equipamentos de suporte (refrigeração). Este artigo reforça a ideia de que a ineficiência é comum em projetos de DATACENTERS projetados há alguns anos. Na verdade, existe um descasamento provocado por instalações construídas em torno do máximo da capacidade e uma carga de TI que está muito longe de consumir a energia disponível.

Este *oversizing* (superdimensionamento) do DATACENTER chega até três vezes a capacidade necessária e era uma prática corriqueira de alguns anos atrás que gerava custos de capital e manutenção em excesso. Este custo pode ser reduzido implementando um arquitetura que possa se adaptar a requisitos de capacidade que mudam com o tempo.

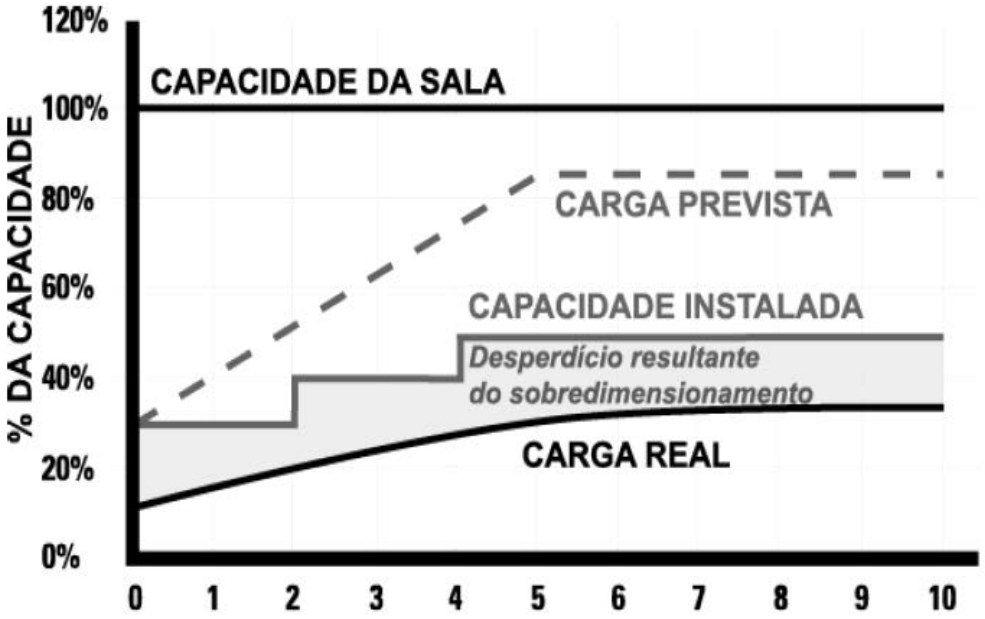
A APC sugere como boa prática a construção do DATACENTER de forma escalonável. A **Figura 4-2 (a) e (b) sugere a diferença entre os projetos**. No caso (a), a capacidade instalada é a própria capacidade da sala, o jeito antigo de fazer o projeto. No caso (b), a ideia é que a capacidade instalada aumenta em incrementos de acordo com o aumento da demanda da carga de TI.

Os fabricantes atualmente, de uma forma geral, correm atrás da opção (b) e estão otimizando os equipamentos utilizados no DATACENTER nos níveis de provimento de energia, refrigeração e carga de TI.



Design de Capacidade de Energia e Requisito sobre o tempo de vida de um DATACENTER.

(a)



Design de Capacidade de Energia e Requisito sobre o tempo de vida de um DATA CENTER.

(b)

Figura 4-2 Design da Capacidade de Energia Típico (a) e Sugerido (b) (Fonte: APC)

A APC sugere que existem cinco fatores chave que contribuem para a ineficiência energética:

- Ineficiência dos equipamentos de energia elétrica.
- Ineficiência dos equipamentos de refrigeração.
- Consumo de energia pela iluminação.
- Superdimensionamento dos sistemas de energia e resfriamento.
- Ineficiências devido à configuração.

Segundo a APC, uma arquitetura ideal, do ponto de vista da eficiência energética, deveria utilizar os seguintes princípios.

- Equipamentos de energia e refrigeração que não são necessários no momento não devem ser energizados.
- O superdimensionamento deve ser reduzido sempre que possível para que os equipamentos possam operar na curva ideal de eficiência.
- Equipamentos de energia, refrigeração e iluminação devem aproveitar as novas tecnologias para reduzir o consumo de energia.
- Subsistemas que devem ser utilizados abaixo da capacidade nominal (devido a redundância) devem ser otimizados para aquela fração de carga e não para a eficiência com carga plena.

4.3. O Green Grid

O Green Grid (*http://www.thegreengrid.org/*) é um consórcio global que foi formado em 2007 por diversas companhias de TI (incluindo Intel, Dell, VMware, AMD), com o objetivo de definir e propagar melhores práticas relacionadas à eficiência no consumo de energia em DATACENTERS. O Green Grid trabalha proximamente ao EPA e ao *Alliance to Save Energy*.

Conforme visto anteriormente, o DATACENTER em operação normalmente é ineficiente do ponto de vista energético. A demanda maior de energia de servidores ultradensos, como os servidores *blade*s, traz um novo problema: a falta de capacidade energética destes antigos DATACENTERS e o surgimento de pontos de alta densidade de energia dentro do DATACENTER que inviabilizam o projeto de refrigeração. Os *RACKS* atuais, por exemplo, podem demandar 25KW de energia devido à densidade atual de potência, o que, alguns anos atrás, era impensável. Estes DATACENTERS foram projetados para máxima funcionalidade e desempenho e, na época, não existia tanta preocupação com o consumo de energia. Além disso, a VIRTUALIZAÇÃO permite mover as cargas de TI dentro do DATACENTER, trazendo mais um componente de complexidade.

Atualmente, alguns fatores guiam os projetos dos novos DATACENTERS:

- Aumento na demanda computacional.
- Aumento na densidade dos equipamentos de TI.
- Utilização de técnicas de VIRTUALIZAÇÃO que permitem mover as cargas de TI dentro do DATACENTER.
- Aumento dos custos de energia.
- Falta de disponibilidade de energia para atender estas novas demandas.

Os custos com energia e refrigeração para servidores ao longo de três anos, por exemplo, já são maiores do que os custos de aquisição, segundo o IDC[5].

Estes fatores tornam as métricas até então utilizadas menos importantes. O caso da métrica DCD (densidade do DATACENTER), usada por anos como a métrica principal do DATACENTER, é emblemático. Esta métrica é excelente quando se quer comparar em termos absolutos o consumo de energia entre um DATACENTER e outro, mas falha quando se quer verificar a eficiência pura do DATACENTER.

DCD = Energia consumida pelos Equipamentos / Área utilizada pelos Equipamentos

Assim, a meta inicial do Green Grid foi a de definir melhores métricas para o consumo de energia do DATACENTER atual, com o objetivo de reduzir o TCO total e, ao mesmo tempo, permitir a competitividade e a adequação para futuras demandas de energia.

5 IDC Worldwide Expense to Power and Cool Installed Server Base, 1996-2010.

No longo prazo, o Green Grid pretende definir uma arquitetura para a definição de políticas e especificações, como também um logotipo vinculado ao compliance e à interoperabilidade de dispositivos para o uso eficiente de energia.

O Green Grid desenvolveu duas novas métricas táticas já bem utilizadas, o PUE (*Power Usage Effectiveness*) e o DCiE (DATACENTER *Efficiency*).

PUE = Energia Total da Instalação / Energia dos Equipamentos de TI

DCiE = 1 / PUE

A energia total da instalação é a energia do medidor que alimenta o DATACENTER. A energia consumida pelos equipamentos de TI é a energia consumida por todos os equipamentos de TI, incluindo KVMs, monitores, estações de gerenciamento. Em muitas instalações não se consideram nem medidores separados para a carga de TI e, portanto, fica difícil medir o PUE.

A **Figura 4-3** ilustra os consumos envolvidos no cálculo do DCiE.

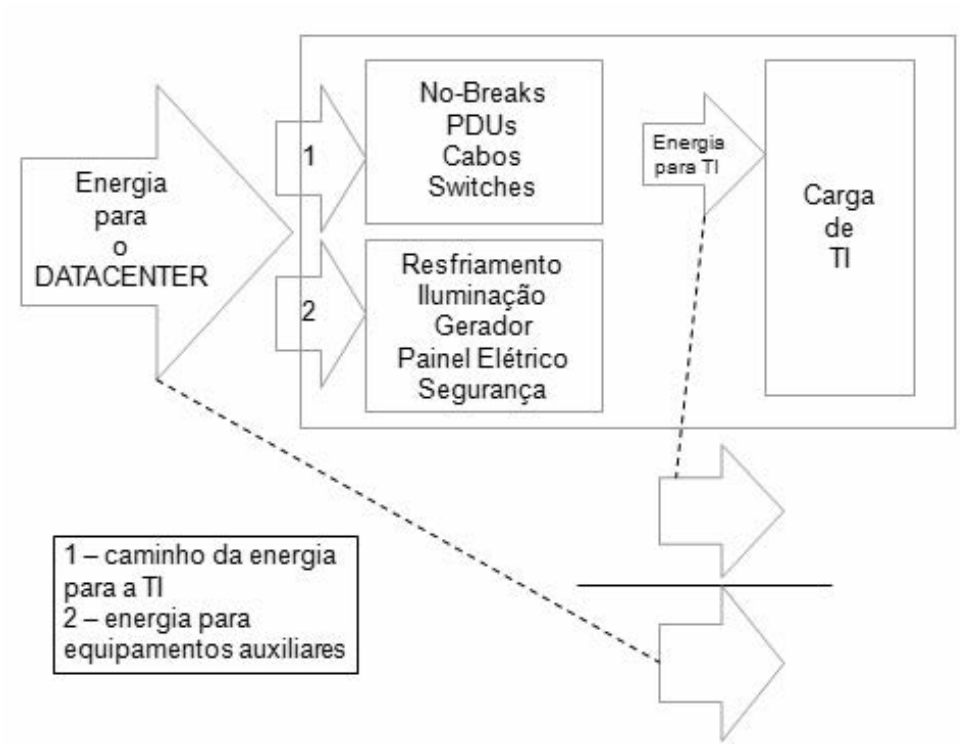


Figura 4-3 Cálculo do DCiE em DATACENTER

Estas duas métricas permitem calcular a real eficiência do DATACENTER, comparar DATACENTERS do ponto de vista do consumo de energia e criar benchmarks, verificar melhoria do consumo ao longo do tempo e oportunidades para realocar energia para novos equipamentos de TI.

Um PUE de 3.0, típico de instalações antigas, indica que o DATACENTER demanda três vezes mais energia do que a necessária para alimentar os equipamentos de TI. Neste caso, se um *RACK* de energia demanda 10KW para alimentar os diversos servidores,

será necessário 30KW na entrada para alimentar todo o DATACENTER. Ou seja, você terá que comprar três vezes mais energia da concessionária. A eficiência é de cerca de 30%. A meta atual para novos DATACENTERS é ter um PUE entre 1,8 e 2,0. O Google fez um grande investimento na melhoria da eficiência dos seus DATACENTERS e recentemente publicou o PUE para dez dos seus principais DATACENTERS, todos abaixo de 1.4.

O Yahoo recentemente construiu um DATACENTER em Nova York que utiliza a tecnologia Yahoo Computing Coop que elimina a necessidade de *chillers*. Segundo o Yahoo, o PUE pode chegar a 1.08. Eu acho difícil. A **Figura 4-4** ilustra o novo DATACENTER do Yahoo.

Figura 4-4 Novo DATACENTER do Yahoo

4.4. Conceito na Prática: APC TradeOff Tools

As ferramentas da *APC , APC TradeOff Tools,* são um conjunto de novos aplicativos baseados na Web com interfaces fáceis de usar, projetadas para utilização nos estágios iniciais de desenvolvimento de conceito e projeto de DATACENTERS. Permitem que gerentes de DATACENTERS testem diversos cenários relacionados à VIRTUALIZAÇÃO, eficiência, dimensionamento de potência, custos de capital e outros aspectos importantes de projeto. O APC TradeOff Tools desmembra as grandes decisões de planejamento dos DATACENTERS em uma série de decisões menores e mais fáceis de gerenciar. O uso dessas ferramentas ajuda a validar, através da modelagem, o design geral de um DATACENTER.

Destacam-se as seguintes ferramentas do conjunto de ferramentas da APC:

- **DATACENTER: Cálculo de Emissão de Carbono**

O DATACENTER Carbon Calculator é uma ferramenta excelente. Ilustra como a localização geográfica, o nível de eficiência e a carga energética do DATACENTER podem impactar as emissões de dióxido de carbono e a conta de luz. Isso oferece aos gestores uma indicação geral de quão "verde" é o DATACENTER e quão "verde" ele poderia ser. A **Figura 4-5** ilustra a tela principal da ferramenta.

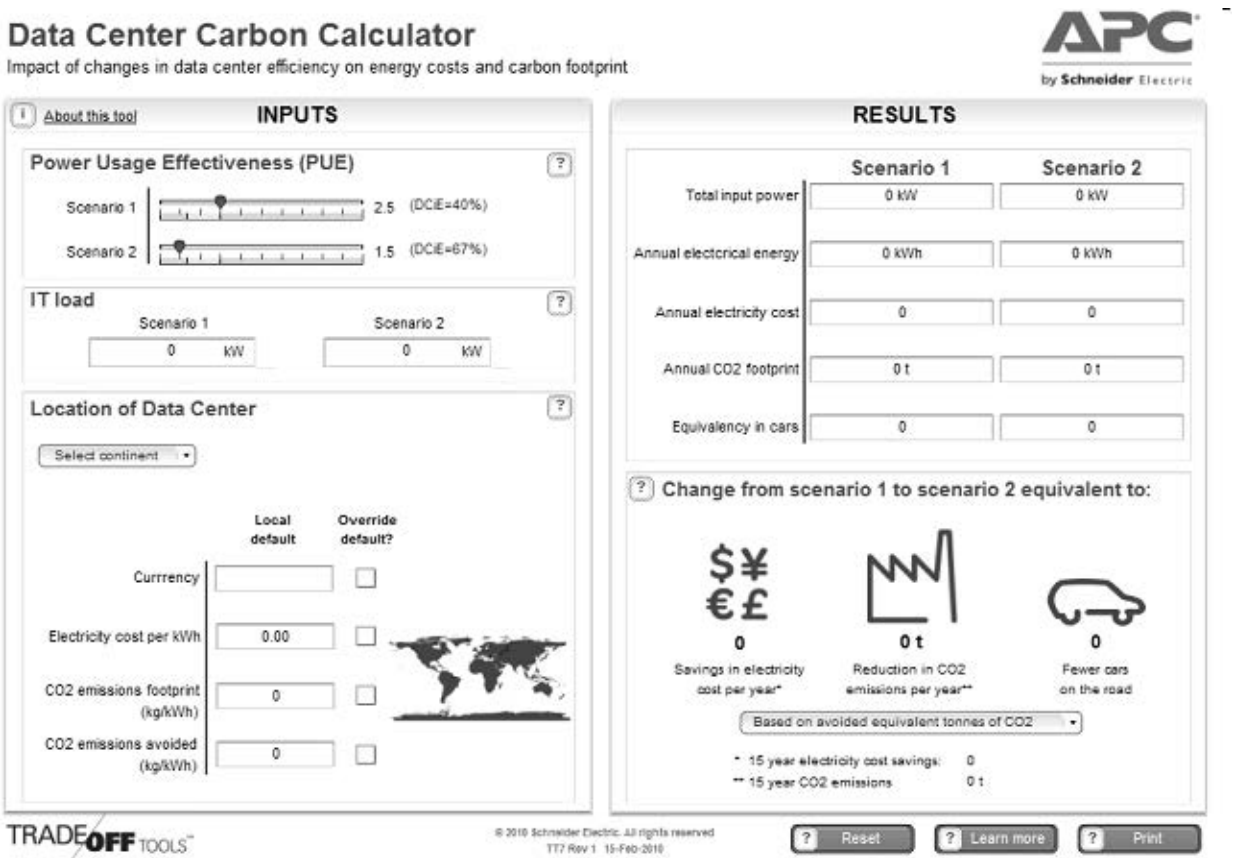


Figura 4-5 Cálculo da Emissão de Carbono

- **DATACENTER: Cálculo de Eficiência Energética**

Essa ferramenta traça o perfil do DATACENTER e calcula o custo de eletricidade e a eficiência resultante com base nas características do ambiente. As empresas podem entender o impacto que cada decisão importante tem na eficiência do DATACENTER. A **Figura 4-6** ilustra a tela da ferramenta.

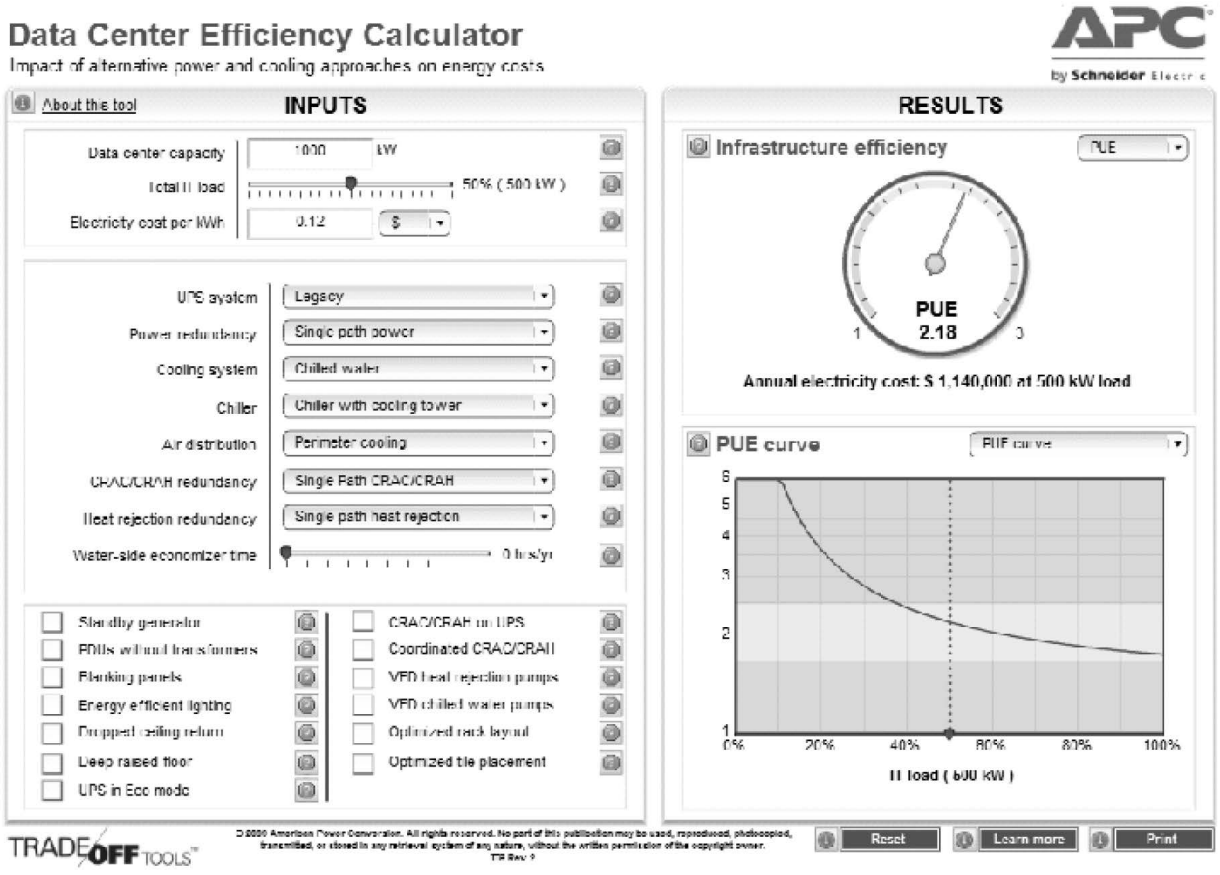


Figura 4-6 Cálculo da Eficiência Energética

- **VIRTUALIZAÇÃO: Cálculo do Custo de Energia**

Abrange características da infraestrutura física e de TI do DATACENTER e calcula a economia de energia resultante da VIRTUALIZAÇÃO de servidores. Isso permite que o usuário teste o impacto que a VIRTUALIZAÇÃO e vários aprimoramentos da infraestrutura física terão no espaço ocupado pelo DATACENTER e no seu consumo de energia. A **Figura 4-7** ilustra a tela da ferramenta.

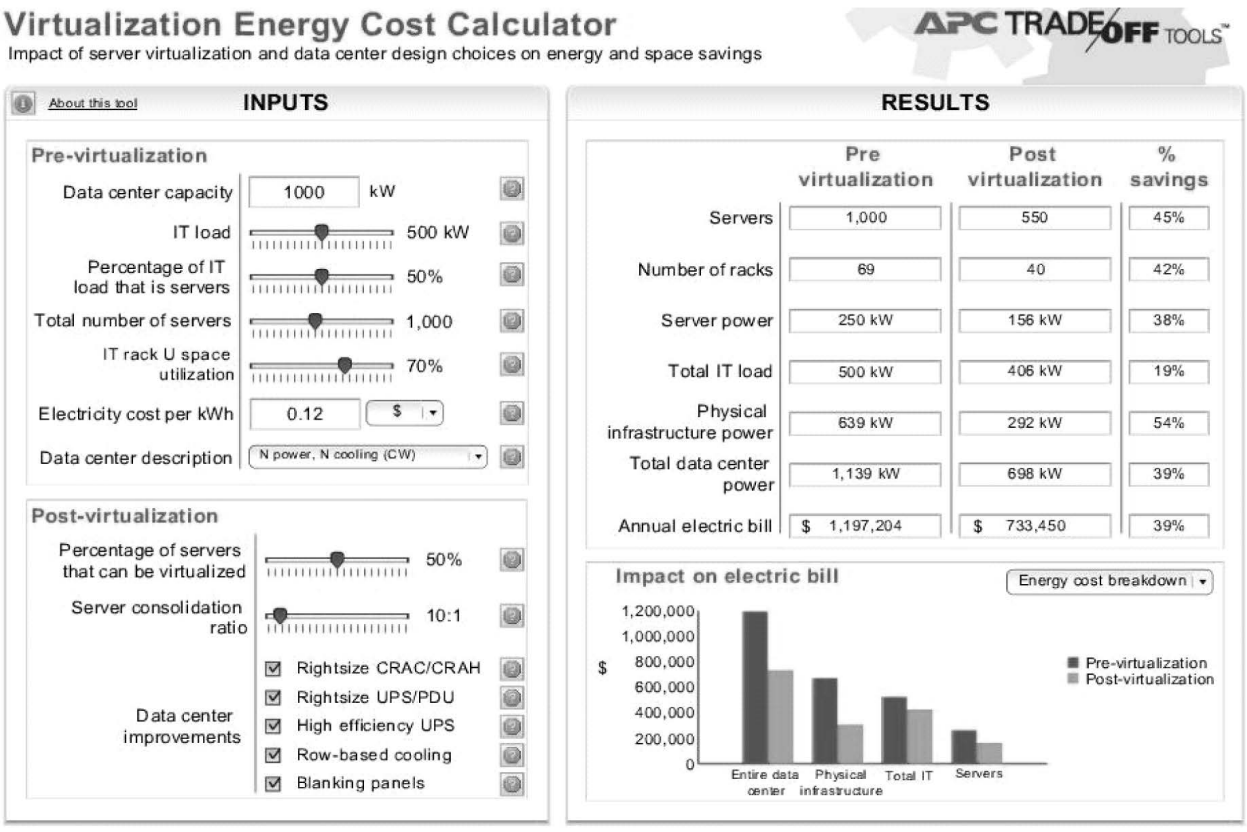


Figura 4-7 Cálculo do Custo de Energia com a VIRTUALIZAÇÃO

O Green Grid vem trabalhando na definição de métricas de longo prazo que possam ser a referência para as instalações de DATACENTER. A DCPE (*DATACENTER Performance Efficiency)* é uma delas e é a natural evolução das métricas anteriores: esta métrica é difícil de determinar, pois trata do real uso da energia pela carga de TI. Diversos esforços estão sendo feitos pela indústria para tornar os equipamentos de TI mais eficientes.

DCPE = Trabalho Útil / Energia Total da Instalação

Outras instituições e ferramentas importantes que foram criadas com foco em TI Verde são descritas abaixo:

- ***DC Pro:*** criado pelo departamento de energia dos Estados Unidos, é um software que ajuda as empresas a diagnosticar quanta energia está sendo utilizada em seu DATACENTER e como ela pode ser economizada.
- ***The Climate Savers Computing Iniciative:*** Organização sem fins lucrativos que tem por objetivo promover o desenvolvimento e a adoção de tecnologias que aumentem a eficiência energética dos computadores.
- ***The Electronic Product Environmental Assessment Tool (EPEAT):*** O EPEAT é uma ferramenta de procurement que auxilia empresas a avaliar, comparar e selecionar desktops, notebooks e monitores com base nos atributos ambientais.

4.5. Requisitos de Energia e Refrigeração para DATACENTERS

O dimensionamento das necessidades de energia de um DATACENTER requer compreensão sobre a quantidade de energia necessária para alimentar o sistema de refrigeração, os no-breaks e a carga de TI. Com a carga de TI prevista, é possível determinar com alguma precisão a necessidade de energia e determinar a capacidade necessária do gerador de reserva. A redundância dos sistemas utilizados é determinada pelos requisitos impostos pelas aplicações de negócio e definem a camada (TIER) do DATACENTER. Os no-breaks podem estar nas configurações N, N+1, 2N. Estimar a energia necessária é um aspecto crucial do projeto. Se estimada para baixo, pode trazer transtornos quando da necessidade de aumento da capacidade do DATACENTER. Se estimada para cima, pode conduzir a custos iniciais elevados. Se o DATACENTER utiliza um prédio, compartilhando as instalações, este cálculo se torna mais complexo. Seguem algumas dicas para realizar este cálculo.

A **Figura 4-8** ilustra um diagrama geral de alimentação do DATACENTER.

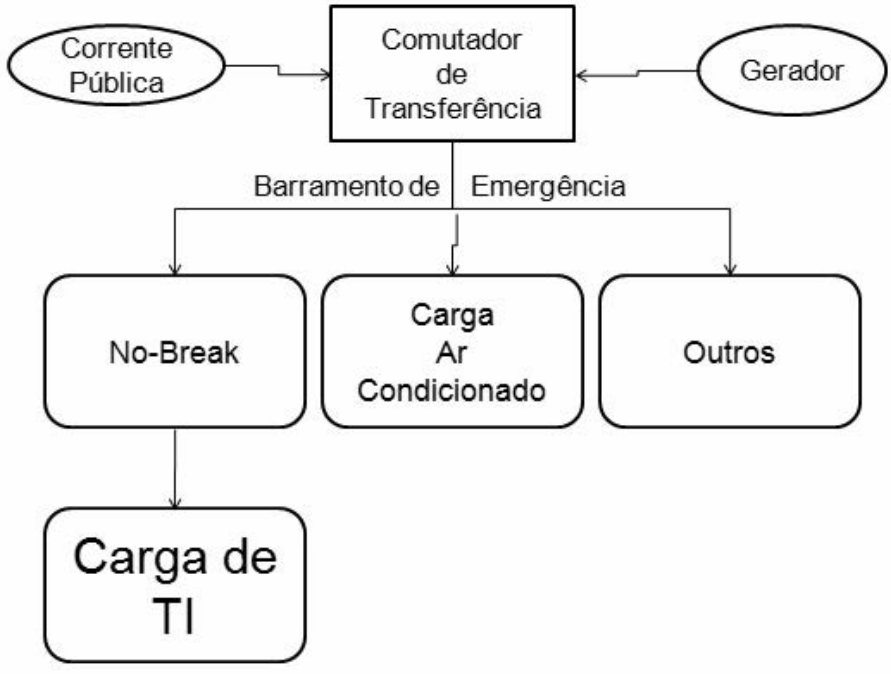


Figura 4-8 Alimentação de um DATACENTER (Fonte: APC)

4.5.1. Carga de TI

A carga de TI deve ser obtida dos dados fornecidos pelos fabricantes referentes ao consumo de energia. Deve-se também fazer uma estimativa de carga futura. Uma possível sequência para o cálculo da carga de TI seria:

- Adicione a potência das cargas previstas. Se houver dificuldade em obter a potência do equipamento, pode-se obter o valor VA de posse da corrente definida na especificação do equipamento e considerando a tensão de alimentação.

- Multiplique o número VA obtido por 0,67, de modo a calcular a potência real em watts.
- Divida o número obtido por 1000 para determinar o nível de carga em KW.
- Deve-se fazer isto também para a carga futura de TI.
- Têm-se assim dois conjuntos de carga de TI: Carga Atual e Carga Futura.

4.5.2. No-Break

Deve-se incluir um fator para a ineficácia do no-break (normalmente a sua eficiência está em torno de 88%) e para a potência adicional necessária para o carregamento da bateria. A potência de carregamento da bateria pode atingir 20% da carga do no-break.

4.5.3. Iluminação

Uma sugestão da APC é prever um valor de 21,5 W por metro quadrado para a iluminação.

4.5.4. Carga da Refrigeração

Todos os equipamentos produzem calor, que deve ser retirado para evitar que a temperatura se eleve a níveis inaceitáveis. A energia térmica produzia pelos equipamentos de TI é, com poucas exceções como roteadores VoIP, igual ao consumo de energia dos equipamentos em watts. Para definir o sistema de refrigeração deve-se conhecer a quantidade de calor produzido pelos equipamentos no espaço a ser utilizado. O calor normalmente é expresso em BTU. É normal tratar a energia produzida em BTU por hora. Para saber o valor em watts precisa-se multiplicar o valor de BTU/Hr por 0.293.

A refrigeração normalmente é baseada em ar ou água. Os sistemas baseados em água variam muito em termos de eficácia. Os sistemas mais comuns, baseados em água, necessitam de 70% do pico da carga suportada. Os sistemas diretos necessitam de 100% do pico da carga suportada.

4.5.5. Dimensionamento do Sistema Elétrico

O dimensionamento do sistema elétrico depende de dois números: carga de TI total e capacidade de refrigeração total. A alimentação elétrica deve suportar estes dois valores junto com a carga de iluminação. O sistema elétrico e o gerador devem ser projetados de acordo com o consumo máximo das cargas.

O sistema elétrico pode então ser calculado da seguinte forma:

- Multiplique a capacidade total em KW por 125%, de modo a respeitar os requisitos das organizações reguladoras.
- Determine a voltagem trifásica de entrada de corrente pública (normalmente 230 Volts CA).

- Utilize a seguinte fórmula para determinar a dimensão do sistema elétrico: Amps = (KW * 1000) / (Volts x 1,73)
- Por fim, pode-se calcular a potência necessária para o gerador.

4.6. Medição do PUE

A APC no relatório técnico 154, cujo título é Medição da Eficiência Elétrica de DATACENTERS, sugere a simplificação da medição do PUE. Para medir a eficiência de um DATACENTER em um ponto operacional em especial, deve-se medir a energia elétrica de entrada total para o DATACENTER e a energia consumida pela carga total da TI.

A medição da energia total de entrada pode ser feita logo depois do comutador de transferência. Este é o caso mais simples, pois o DATACENTER tem uma entrada exclusiva de energia.

O mais complicado é medir a energia consumida pela carga de TI. Se a carga for um único equipamento gigantesco, a energia elétrica da carga da TI será uma única medição na conexão elétrica do equipamento. Seriam necessárias apenas duas medições neste caso hipotético. Infelizmente, tal situação ideal nunca ocorre.

A maioria dos DATACENTERS faz parte de prédios multiuso com outras cargas, e todos os DATACENTERS são constituídos de um conjunto de equipamentos de TI – possivelmente milhares – com circuitos elétricos separados. A ideia é utilizar o conceito de carga agregada de TI simplificando a medição de consumo de diversas cargas de TI. A sugestão dada pela APC é medir a energia consumida pela carga de TI logo depois do no-break, mas considerar as perdas com as unidades de distribuição de energia (PDUs) introduzidas, conforme ilustra a **Figura 4-9**.

Existem duas principais abordagens para tentar aumentar a eficiência do DATACENTER que podem e devem ser utilizadas conjuntamente. A primeira trata de melhorar a eficiência da infraestrutura de apoio, que são os equipamentos que transformam a energia e refrigeram o ambiente. A segunda trata a eficiência no nível da carga de TI. Os principais instrumentos para atender a estas abordagens são o planejamento da capacidade instalada e a VIRTUALIZAÇÃO.

- **Planejar a capacidade instalada**: dimensionamento correto da infraestrutura para a carga de TI esperada. Envolve conhecer melhor o consumo dos diversos equipamentos utilizados e fazer um projeto que utilize componentes escaláveis. Os projetos de DATACENTER de alguns anos atrás consideravam a capacidade instalada muito próxima da capacidade da sala. A nova proposta para a capacidade instalada é ter equipamentos de energia e refrigeração que possam crescer (escalar) com o passar do tempo. Esta técnica possibilita uma grande redução de consumo de energia do DATACENTER ao longo dos anos, mas muitas vezes é de difícil execução. Neste novo formato a capacidade instalada começa num patamar muito menor e cresce ao longo dos anos. É evidente que esta forma só é possível se os componentes (no-breaks, equipamentos de refrigeração, etc.) utilizados forem escaláveis.

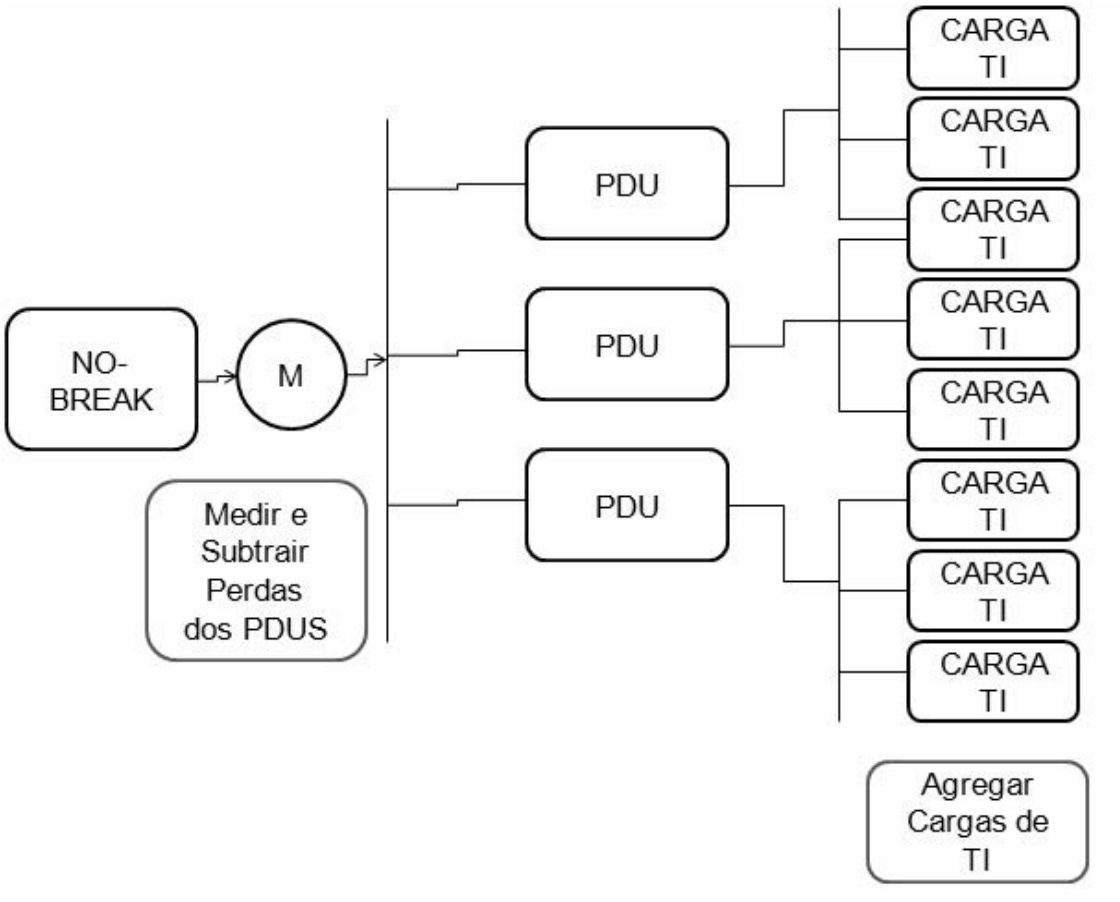


Figura 4-9 Medição do consumo da carga de TI (Fonte: APC)

- **Utilizar a VIRTUALIZAÇÃO**: a VIRTUALIZAÇÃO traz inúmeras vantagens ao ambiente de DATACENTER. A VIRTUALIZAÇÃO de servidores que permite o processamento em um único servidor de cargas de trabalho múltiplas, cada uma com seu nível de serviço, é uma aliada na redução de calor, considerando que um servidor a plena carga ou a 15% consome praticamente a mesma energia. A possibilidade de alocar dinamicamente estas cargas de trabalho para atender aos requisitos do negócio meio que, por si só, já justificam a VIRTUALIZAÇÃO. Este conceito, conhecido como mobilidade da carga de trabalho, pode ser estendido para o gerenciamento térmico e de energia do DATACENTER. Em períodos de baixa utilização, parte dos servidores de um DATACENTER pode ser transicionada para um estágio de baixo consumo e, desde que a infraestrutura de energia e refrigeração possa estar integrada, esta ideia pode reduzir o consumo de energia de maneira interessante.

Outras medidas relevantes são fazer escolhas inteligentes sobre a localização de servidores, medir e monitorar a eficiência das instalações e fazer uso de melhores práticas. Também conta substituir os equipamentos legados por sistemas que oferecem alto desempenho por watt. Equipamentos antigos muitas vezes são ineficientes e, em certas situações, não vale a pena esperar o final do ciclo do seu uso.

A APC sugere que, para vencer os desafios impostos pela VIRTUALIZAÇÃO ao projeto de DATACENTERS, sejam adotadas as seguintes medidas:

- Para cargas de altas densidades dinâmicas utilizar resfriamento baseado em colunas de RACKS. Em instalações convencionais de DATACENTER (corredor quente/corredor frio) fazer a retirada dos CRACs (Computer Room Air Conditioning) e fazer a refrigeração no próprio RACK. A solução de refrigerção no RACK varia a velocidade do ventilador de acordo com a temperatura.
- Para sistemas de energia e refrigeração superdimensionados utilizar equipamentos de energia e refrigeração escaláveis.
- Utilizar ferramentas de gerenciamento de capacidade (*capacity management tools*) para saber se a capacidade atende à demanda no nível da fila, RACK e servidor.

Outra alternativa sugerida pela APC é criar uma zona de alta densidade para equipamentos do tipo blades com sistema exclusivo de refrigeração.

Deve ficar claro que a eficiência no consumo da carga de TI é melhorada com a VIRTUALIZAÇÃO, mas a eficiência global do DC é otimizada se o DCiE for otimizado também. Ou seja, os equipamentos de energia e refrigeração precisam também estar adequados ao consumo menor imposto pela VIRTUALIZAÇÃO, e o projeto deve contemplar a VIRTUALIZAÇÃO.

4.7. Impactos da VIRTUALIZAÇÃO no Green DATACENTER

A VIRTUALIZAÇÃO é a tecnologia que permitiu que o DATACENTER pudesse alocar recursos em função da demanda. A otimização propiciada pela VIRTUALIZAÇÃO trouxe inúmeras vantagens, incluindo a redução do consumo de energia.

A otimização do uso dos recursos propiciada pela VIRTUALIZAÇÃO possibilita, junto com outras medidas, reduzir o consumo de energia do DATACENTER.

4.8. Referências Bibliográficas

Best Practices for Unlocking your Hidden DATACENTER, Dell Power Solution, fev. 2008.

Calcular requisitos de potência totais em centros de dados, AT 3 revisão 1, APC.

Cálculo dos requisitos de refrigeração total para central de dados, APC Relatório Interno 25, APC.

DATACENTER Efficiency in the Scalable Enterprise, Dell Power Solutions, 2007.

Deploying High-Density Zones in a Low-Density DATACENTER. APC White Paper #134.

Evitando despesas por obter um DATACENTER e Sala de Infraestrutura de Rede maior que o necessário. APC White Paper #37.

Green Grid Metrics: Describing DATACENTER Power Efficiency, The Green Grid, 2007.

Medição da Eficiência Elétrica de DATACENTERS, APC White paper #154.

O DATACENTER verde, IBM, 2007.

The Green Grid Oportunity: decreasing DATACENTER and other IT energy usage patterns, The Green Grid, Fev. 2007.

Uma arquitetura aperfeiçoada para DATACENTERS de alta densidade e alta eficiência. APC White Paper #126.

Veras, Manoel. DATACENTER: Componente Central da Infraestrutura de TI, Brasport, 2009.

Verdi, Fábio Luciano; Rothenberg, Christian & Pasquini, Rafael, Magalhães, Maurício Ferreira, Novas Arquiteturas de DATACENTER para Cloud Computing, 2010.

Virtualization: Optimized Power and Cooling to Maximize Benefits. APC White Paper #118.

PARTE II: TECNOLOGIA DE VIRTUALIZAÇÃO

5. Conceitos Centrais

5.1. Introdução

A VIRTUALIZAÇÃO é a tecnologia central de um DATACENTER e essencialmente transforma, obedecidas certas condições, um servidor físico em vários servidores virtuais. Os impactos são muitos e os benefícios também.

A Forrester Consulting fez uma pesquisa[6] que entrevistou vice-presidentes e diretores do nível C em vinte e nove grandes corporações e constatou que a VIRTUALIZAÇÃO de servidores trouxe os seguintes benefícios para as suas organizações:

- TI se tornou mais eficiente.
- A organização passou a ter um Time-to-Market mais adequado.
- Os serviços de TI se tornaram mais previsíveis.

A Forrester observou que as organizações deixam de ter maiores ganhos devido ao fato de terem pequenas taxas iniciais de consolidação que variam de 10% a 30% do parque de servidores e poucas iniciativas em direção a novos projetos de VIRTUALIZAÇÃO.

A pesquisa também verificou que as taxas de consolidação já são maiores do que 10:1 para quase 50% das organizações pesquisadas e o ROI acontece em menos de doze meses para boa parte delas. O motivo principal apontado pelas organizações que fizeram parte da pesquisa para fazer a VIRTUALIZAÇÃO de servidores foi a redução de custos de hardware, em primeiro, e a melhoria da condição de recuperação a desastres e da continuidade do negócio, em segundo.

Os servidores virtuais criados com a VIRTUALIZAÇÃO oferecem um ambiente similar ao de um servidor físico e otimizam o uso de recursos, tornando as aplicações independentes do hardware. Transforma-se assim um ambiente baseado em servidores físicos em um ambiente baseado em servidores virtuais ou máquinas virtuais. A ideia central é sair de um ambiente 1:1 (uma aplicação para um servidor físico) para N:1 (N aplicações para um servidor físico).

---

6 FORRESTER: The Business value of VIRTUALIZATION.

A VIRTUALIZAÇÃO traz normalmente outras vantagens como a redução do espaço físico utilizado pela infraestrutura de TI e, como permite distribuir a carga de trabalho entre um conjunto de servidores, otimiza o uso da infraestrutura.

A VIRTUALIZAÇÃO também está alinhada ao conceito de TI verde, já que permite obter uma economia significativa de energia no DATACENTER.

A redução dos custos é o grande fator motivador da VIRTUALIZAÇÃO. Pesquisas do IDC dão conta que só 15% da capacidade dos servidores é utilizada nas empresas que não usam a VIRTUALIZAÇÃO. Os 85% restantes estão ociosos. Com a VIRTUALIZAÇÃO, este número de 15% pode subir para pelo menos 60%.

A evolução da arquitetura dos processadores, a proliferação de redes mais velozes e a introdução de unidades de armazenamento com grande escalabilidade e baixo custo, além do desenvolvimento de softwares específicos, fizeram a VIRTUALIZAÇÃO também passar a ser utilizada por pequenas instalações de TI. Ou seja, a VIRTUALIZAÇÃO também é uma realidade para pequenas empresas.

5.2. Workload e Throughput

Dois conceitos fundamentais em VIRTUALIZAÇÃO são os de WORKLOAD E THROUGHPUT.

A WORKLOAD constitui-se dos dados a serem processados e as instruções a serem executadas sobre estes dados e define a qualidade de serviço percebida pelo usuário na ponta. Logicamente, a WORKLOAD varia de acordo com a aplicação a ser processada. Assim, ela é definida pela demanda da aplicação (muitas vezes a demanda do banco de dados gerencial, o SGBD). Combina solicitações como transações online, batch jobs, consultas ad hoc, data warehousing, consultas analíticas e comandos dirigidos à aplicação. Sobre a WORKLOAD é correto afirmar:

- Pode variar drasticamente de dia para dia, hora para hora, minuto para minuto, segundo para segundo.
- Algumas vezes é previsível, mas muitas vezes é imprevisível.
- Tem grande impacto no desempenho da aplicação.

O THROUGHPUT define a capacidade do hardware/software para processar dados.

- É composto de velocidade de I/O, velocidade de CPU, capacidades de paralelismo e eficiência do SO.
- O núcleo do sistema operacional, o espaço em disco, os controladores de cache e o microcódigo são exemplos dos recursos a serem avaliados para definir o THROUGHPUT.

A WORKLOAD e o THROUGHPUT são as unidades básicas utilizadas quando do dimensionamento da infraestrutura virtualizada.

5.3. Conceito

A VIRTUALIZAÇÃO pode ser conceituada de duas principais formas:

- É o particionamento de um servidor físico em vários servidores lógicos. Na **Figura 5-1** cinco servidores são substituídos por um servidor virtual e, portanto, a taxa de consolidação é de 5:1.

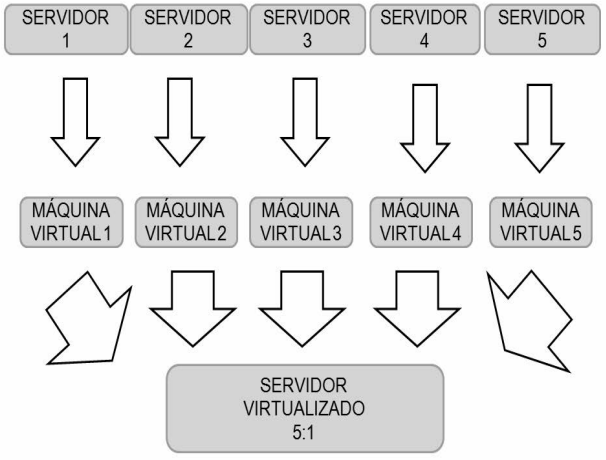


Figura 5-1 O que é a VIRTUALIZAÇÃO

- É uma camada de abstração entre o hardware e o software que protege o acesso direto do software aos recursos físicos do hardware. A VIRTUALIZAÇÃO permite que a camada de software (aplicações e sistema operacional) seja isolada da camada de hardware. Normalmente a VIRTUALIZAÇÃO é implementada por um software. A **Figura 5-2** ilustra o conceito.

A VIRTUALIZAÇÃO simplifica o gerenciamento e permite flexibilizar e ampliar o poder de processamento. Funcionalidades contidas nos softwares de VIRTUALIZAÇÃO também permitem melhorar a disponibilidade e a recuperação de desastres de ambientes de TI de uma maneira mais simples e com menor custo quando comparado a formas tradicionais.

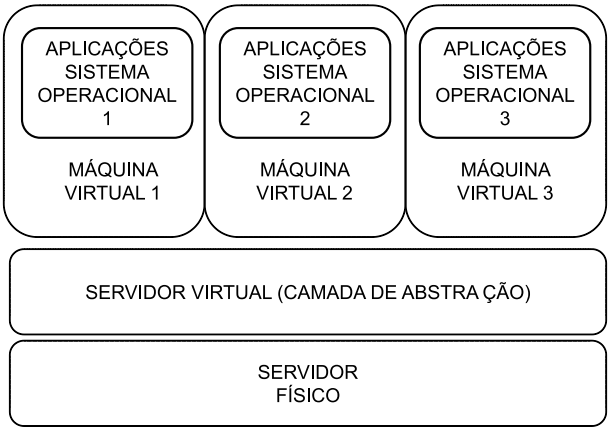


Figura 5-2 Arquitetura da VIRTUALIZAÇÃO

A camada de VIRTUALIZAÇÃO entrega para o sistema operacional convidado um conjunto de instruções de máquinas equivalente ao processador físico. A camada de VIRTUALIZAÇÃO de servidores mais conhecida é o HYPERVISOR ou MONITOR DE MÁQUINA VIRTUAL (*Virtual Machine Monitor – VMM*). O servidor físico virtualizado pode então rodar vários servidores virtuais chamados de máquinas virtuais (*virtual machines – VMs*).

Com a VIRTUALIZAÇÃO, cada VM utiliza um sistema operacional e suas respectivas aplicações. Com a VIRTUALIZAÇÃO, diversas VMs podem coexistir no mesmo servidor físico.

A **Figura 5-3** ilustra a VIRTUALIZAÇÃO com HYPERVISOR.

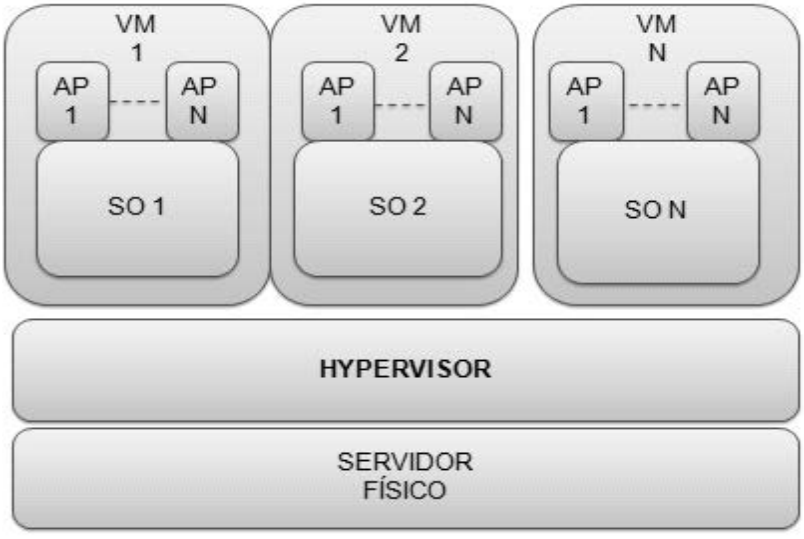


Figura 5-3 Virtualização com HYPERVISOR

5.4. Máquina Virtual

Uma máquina virtual é um CONTAINER de software totalmente isolado e capaz de executar sistemas operacionais e aplicações próprias como se fosse um servidor físico. Uma máquina virtual se comporta exatamente como um servidor físico e tem CPU, memória RAM, disco rígido e NIC (*Network Interface Card*, placa de interface de rede) virtuais baseados em software. A **Figura 5-4** ilustra a ideia central da máquina virtual.

Uma máquina virtual (VM) pode também ser definida como uma duplicata eficiente e isolada de uma máquina real. Os recursos de processamento, memória e outros são virtualizados.

A diferença entre uma máquina virtual e um servidor físico não pode ser notada por um sistema operacional, muito menos por aplicações ou outros computadores na rede. A máquina virtual é constituída por software e não contém componentes de hardware. Como resultado, as máquinas virtuais oferecem a vantagem da flexibilidade em comparação com o hardware físico.

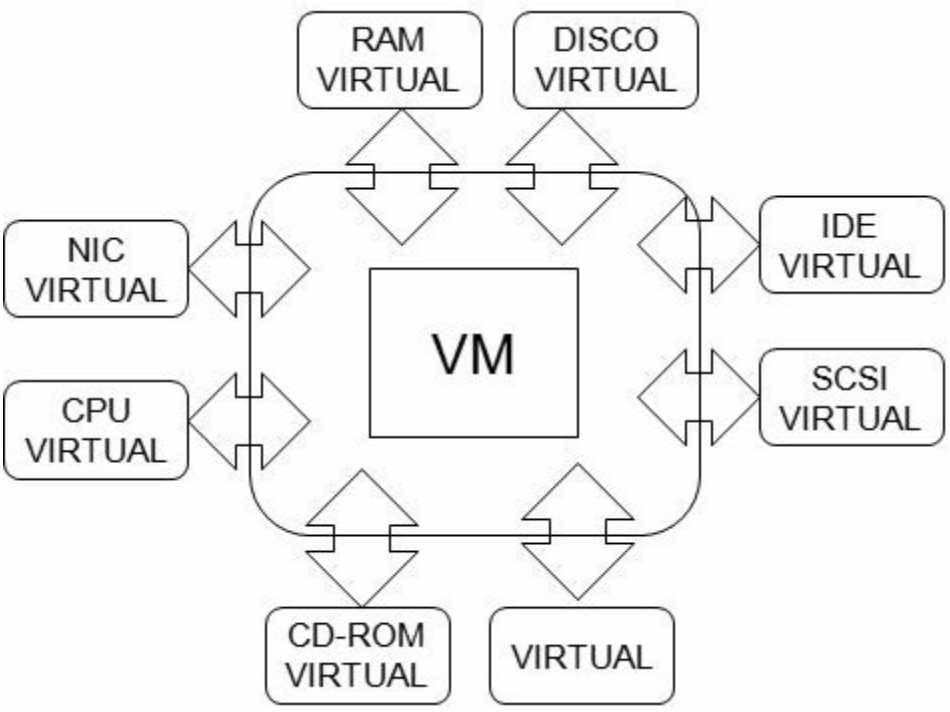


Figura 5-4 Máquina Virtual – *Virtual Machine* – VM

5.5. Histórico

Pode-se dizer que a ideia da VIRTUALIZAÇÃO iniciou com a publicação do artigo *Time Sharing Processing in Large Fast Computers,* por Christopher Strachey, cientista da computação, na Conferência Internacional de Processamento de Informação realizada em Nova York em 1959. Sua publicação tratou do uso da multiprogramação em tempo compartilhado e estabeleceu um novo conceito de utilização de máquinas de grande porte. Essas máquinas de grande porte poderiam agora utilizar melhor os recursos de hardware.

Baseado no trabalho inicial de Strachey, o MIT desenvolveu o padrão CTSS (*Compatible Time Sharing System*), utilizado como referência por diversas empresas que fabricavam máquinas de grande porte.

A IBM posteriormente introduziu o multiprocessamento nos mainframes, baseado na evolução do padrão CTTS, que permitiu que várias unidades de processamento trabalhassem como uma só, antecipando o conceito de VIRTUALIZAÇÃO. Estes mainframes introduziram o conceito de memória virtual (também chamado de storage virtual) como parte do sistema operacional. Esta opção possibilitou a abstração e o mapeamento da memória real para memória virtual e a especificação de partições ou espaços de endereçamento que eram utilizadas por programas diferentes. Surgiam as primeiras formas de fazer a VIRTUALIZAÇÃO.

A VIRTUALIZAÇÃO inicialmente utilizou o conceito de máquina virtual de processo. Uma máquina virtual de processo é uma aplicação que é executada em um sistema

operacional A e que emula o comportamento de um sistema operacional B. Assim, aplicações desenvolvidas para o sistema operacional B podem rodar em um sistema operacional A. É importante salientar que essa técnica de implementação permite que binários de um processador sejam interpretados e substituídos por código equivalente de outro processador. Portanto, além de emular o sistema operacional, é possível emular processadores. As desvantagens dessa técnica são basicamente duas: desempenho e desperdício de capacidade do hardware físico. O desempenho é sacrificado, já que há uma tradução de um sistema para outro. O desperdício da capacidade física do hardware vem do fato que as máquinas virtuais de processo oferecem dispositivos de I/O genéricos e simples e este fato sobrecarrega o sistema operacional.

Os monitores de máquinas virtuais VMM (*Virtual Monitor Machine*) ou HYPERVISORS surgiram para resolver as desvantagens das máquinas virtuais de processos. Eles são implementados como uma camada de software entre o hardware e o sistema operacional, oferecendo uma máquina virtual para o sistema operacional. São mais eficientes por tratar de outra forma os dispositivos de I/O.

O sistema operacional VM (*virtual machine*), da IBM, surgiu baseado no conceito de HYPERVISOR. O HYPERVISOR é esta camada de software que permite que vários sistemas operacionais diferentes rodem de maneira isolada em um único hardware. O sistema operacional IBM VM/370, baseado na VIRTUALIZAÇÃO, foi bastante utilizado para migração de um mainframe para outro ou de um sistema operacional para outro, desde que ele permitisse que os sistemas novo e velho rodassem juntamente com a supervisão do HYPERVISOR.

Com o crescimento do poder de processamento dos servidores x86 e o aumento da sua confiabilidade, e também a disponibilidade de redes de longa distância com qualidade aceitável, ficou evidente que a VIRTUALIZAÇÃO era um instrumento importante para a melhoria do TCO dos DATACENTERS. Ela de alguma forma permitiria uma nova centralização dos servidores, agora baseados em baixa plataforma, na medida em que seria possível rodar várias aplicações em um único servidor e melhoraria o gerenciamento. Mas existia uma dificuldade tecnológica para a implementação da VIRTUALIZAÇÃO em servidores x86 que só foi resolvida no final da década de 90.

O marco inicial desta nova era foi o surgimento da VMware, em 1998, criada por Diana Greene e Mendel Rosenblun. A VMware desenvolveu o primeiro HYPERVISOR que permitiu a VIRTUALIZAÇÃO de servidores em plataformas baixas x86. Já existia outra empresa desde 1996, a Connectix, fundada em 1988, que tratava de VIRTUALIZAÇÃO, mas em ambiente Mac. Posteriormente, a Microsoft adquiriu a Connectix (2003) e a EMC adquiriu a VMware (2004). A Microsoft lançou o Microsoft Virtual Server 2005, seu primeiro produto com foco na VIRTUALIZAÇÃO de servidores, já em 2004.

Com o desenvolvimento da tecnologia de VIRTUALIZAÇÃO em servidores x86 e a adaptação tecnológica realizada por diversos fabricantes de processadores, hardware e sistemas operacionais, foram desenvolvidas outras funcionalidades que ajudaram a justificar o uso e permitiram melhorar o TCO e o nível de serviços fornecidos pelo DATACENTER virtualizado.

A técnica de VIRTUALIZAÇÃO atendeu a uma demanda antiga da área de TI provocada pelo mau uso de servidores x86, que, muitas vezes, rodam uma única aplicação utilizando um mínimo de recursos. Seria complicado, para efeitos de segurança e disponibilidade, fazer rodar mais de uma aplicação em um mesmo servidor físico sem uma camada que isolasse a aplicação do hardware. A VIRTUALIZAÇÃO resolveu a dificuldade.

Conforme dito anteriormente, um aspecto importante de qualquer projeto de VIRTUALIZAÇÃO é a possibilidade de redução do consumo de energia dos servidores e da consequente energia gasta com a refrigeração. O fato de otimizar o uso de recursos em um servidor promove a economia de energia e refrigeração, pois sabe-se que um servidor a plena carga e um servidor sem carga consomem energia de maneira muito próxima[7].

A VIRTUALIZAÇÃO da forma que é feita hoje é uma revolução na área de TI. Isso é confirmado por análises econômicas conduzidas pela Gartner, pela criação de associações como a EMA (*Enterprise Management Association*) e pelo crescimento exponencial de empresas como a VMware. O IDC prevê que o crescimento do mercado e dos investimentos em VIRTUALIZAÇÃO subirá de U$6.5 bilhões para U$15 bilhões em 2011.

5.6. Benefícios

Os indutores da VIRTUALIZAÇÃO são as novas demandas oriundas do diretor de TI ou CIO (*Chief Executive Oficcer*), do gerente de infraestrutura e tendências de mercado. Os CIOs priorizam usar a TI para o crescimento dos negócios, consolidar as operações e introduzir flexibilidade, provar o valor do negócio de TI. Os gerentes de infraestrutura querem reduzir o número de servidores, melhorar a disponibilidade dos sistemas, mitigar os riscos da mudança, reduzir o tempo para realizar o *deployment* de servidores e fazer mais com menos. O mercado se encarrega de dar os sinais de que uma nova tendência está chegando.

Essencialmente, é necessário entender os objetivos do negócio considerando que CIOs e gerentes de infraestrutura possuem visões diferentes das necessidades, mas que convergem em um ponto quando o tema é transformar a situação atual da infraestrutura de TI da organização.

As justificativas para novos projetos que contemplem a VIRTUALIZAÇÃO passam por dificuldades atuais em provisionar novos servidores, dificuldades operacionais, consumo de energia, mau uso dos recursos, falta de estratégia para alta disponibilidade e recuperação de desastres, dificuldades com as aplicações legadas e novos hardwares.

Os desafios são diversos em um projeto de VIRTUALIZAÇÃO, dentre eles o principal é o de alterar a infraestrutura existente, considerando que a organização continua a funcionar. Fazer a mudança para o novo ambiente de uma forma que haja o mínimo de paralisação das operações com a certa limitação de recursos humanos e físicos é um desafio.

---

7 Best Practices for Unlocking your Hidden DATACENTER, Dell Power Solution, fev. 2008.

As organizações avançam na busca de soluções não proprietárias que simplifiquem o ambiente de infraestrutura de TI. Cerca de 90% dos servidores vendidos atualmente são baseados em plataforma x86, portanto esta é uma tendência de mercado. A VIRTUALIZAÇÃO permite utilizar melhor os recursos de hardware, incluindo aperfeiçoar o uso dos recursos de processador e a memória dos servidores.

Novos negócios exigem uma infraestrutura de TI ágil e flexível para suportar a dinâmica dos processos. O uso da VIRTUALIZAÇÃO pode trazer grandes benefícios para a organização com a adequação da infraestrutura de TI ao negócio, mas requer planejamento e aquisição de novos recursos.

Os prováveis benefícios obtidos com a VIRTUALIZAÇÃO são:

- **Redução do TCO**: O TCO pode ser reduzido com o uso da técnica de VIRTUALIZAÇÃO. Os fabricantes disponibilizam ferramentas que permitem o cálculo do TCO considerando a comparação de uma infraestrutura de TI com e sem VIRTUALIZAÇÃO. Em geral é simples de justificar um projeto de VIRTUALIZAÇÃO utilizando a abordagem de TCO, tanto para a atualização da infraestrutura física já existente como também para a construção de uma nova infraestrutura. A redução do TCO tem a ver com os seguintes aspectos:
  - **Redução do uso do espaço físico:** a utilização da VIRTUALIZAÇÃO permite a redução do espaço físico, na medida em que considera a utilização de menos servidores como solução. Também a consolidação das estruturas de storage e backup, quase sempre contempladas num projeto de VIRTUALIZAÇÃO de servidores, acaba reduzindo a utilização do espaço como um todo.
  - **Redução do consumo de energia:** quase sempre junto com a consolidação física vem a redução do consumo de energia. Servidores são os responsáveis pelo maior consumo de energia entre os equipamentos de TI, e a consolidação acaba por reduzir o consumo de energia.
  - **Isolamento dos ambientes de testes, desenvolvimento e produção**: em muitas instalações construir ambientes físicos diferentes para os ambientes de teste, desenvolvimento e produção pode ser muito caro. A utilização da VIRTUALIZAÇÃO permite otimizar o uso dos recursos, pois permite que estes ambientes existam de maneira completamente isolada, mesmo estando em poucos servidores físicos.
  - **Flexibilidade na criação de novas máquinas virtuais**: as máquinas virtuais podem ser criadas de forma automática em servidores já existentes. Na prática, a demanda por um novo servidor físico que dependeria de aprovação, compra, entrega, etc. pode ser atendida por uma máquina virtual pronta para rodar.
  - **Padronização das plataformas**: na medida em que o HYPERVISOR passa a ser o elemento central do servidor virtualizado, todo o esforço de padronização de plataforma fica simplificado, pois a relação com o hardware se dá através dele. Diferentes sistemas operacionais podem coexistir sobre a arbitração do HYPERVISOR.

- **Gerenciamento centralizado**: o gerenciamento das máquinas virtuais fica centralizado em uma única ferramenta com única interface, reduzindo os custos operacionais de gerenciamento e promovendo a simplificação do ambiente.
- **Simplifica a implantação de técnicas de alta disponibilidade e recuperação de desastres**: a implantação de técnicas de alta disponibilidade, como clusters de servidores e o uso de tecnologias de replicação para suportar a recuperação a desastres, pode ser simplificada com o uso da VIRTUALIZAÇÃO. A VIRTUALIZAÇÃO permite a utilização do recurso de alta disponibilidade independentemente da técnica de cluster e facilita a criação do site secundário otimizando os recursos alocados para o segundo site. Além disso, permite automatizar os processos de recuperação de desastres com a fácil integração promovida com técnicas de replicação do storage.

- **Viabiliza a CLOUD COMPUTING e o DATACENTER DINÂMICO**: a VIRTUALIZAÇÃO é o componente central do DATACENTER DINÂMICO, que, por sua vez, viabiliza a CLOUD COMPUTING. A CLOUD COMPUTING e o DATACENTER DINÂMICO se viabilizam na medida em que as soluções de VIRTUALIZAÇÃO avancem nos aspectos referentes a balanceamento de carga dinâmica, recuperação de falhas, segurança e interoperabilidade de sistemas diferentes.

5.7. Principais Fornecedores

Os principais fornecedores de software de VIRTUALIZAÇÃO para servidores empresariais são a VMware, a Microsoft e a Citrix.

- **VMware**: VMware ESXi, ESX e VMware vSphere.
- **Microsoft**: Hyper-V e Microsoft Windows Server com Hyper-V.
- **Citrix**: Xen Server; Citrix Essentials for Hyper-V e Citrix Essentials for Xen Server.

Existem outros fornecedores como a Red Hat, Parallels, Oracle e Novell.

Estas empresas comercializam os softwares de VIRTUALIZAÇÃO de diversas formas e com diversas funcionalidades. Pesquisa realizada pelo Enterprise Strategy Group com 365 respondentes de grandes organizações nos Estados Unidos, em 2008, indicou que todas as organizações pesquisadas já utilizam a VIRTUALIZAÇÃO de alguma forma. A maioria das organizações utiliza o software da VMware como principal ferramenta de VIRTUALIZAÇÃO, seguido da Microsoft e da Citrix.

A **Figura 5-5** ilustra o quadrante mágico do GARTNER[8] para a VIRTUALIZAÇÃO baseada na plataforma x86 e demonstra a superioridade dos produtos da VMware, Microsoft e Citrix.

8 Magic Quadrant for x86 Server Virtualization Infrastrucuture, GARTNET, 2010.

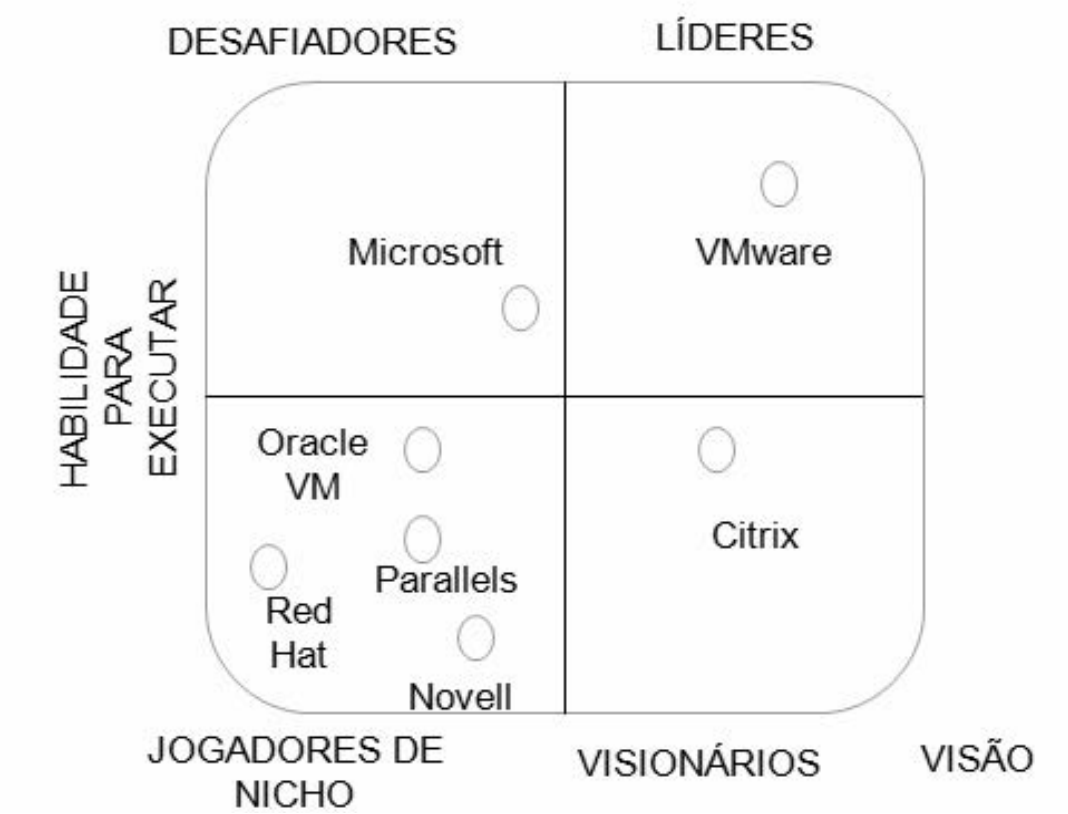


Figura 5-5 Quadrante Mágico do Gartner para VIRTUALIZAÇÃO x86

5.8. Desempenho e Benchmarks

A VIRTUALIZAÇÃO é uma técnica que permite que um sistema computacional (nativo) execute vários outros denominados de máquinas virtuais (convidados). Cada máquina virtual, por si só, é um ambiente de execução isolado e independente das demais. Com isso, cada máquina virtual pode ter seu próprio sistema operacional, aplicativos e serviços de rede. O sistema operacional nativo pode ser diferente daquele utilizado pelo convidado.

Como a VIRTUALIZAÇÃO consiste basicamente em pôr uma camada de software a mais em um sistema computacional, a questão sobre quanto isso afeta o desempenho final é imediata. Estudos feitos pela VMware e pela XenSource apontam para um queda de desempenho causada pela presença do HYPERVISOR, em geral, entre 2% e 10%, com algumas situações impondo perdas maiores. Cabe ressaltar que esses resultados foram obtidos usando benchmarks genéricos.

Os fabricantes envolvidos com a VIRTUALIZAÇÃO, e não são poucos, estão fazendo muitos esforços no sentido de reduzir a queda de desempenho acarretada pela camada de VIRTUALIZAÇÃO.

Para que haja uma forma padrão, e isenta, de avaliar o desempenho das diversas plataformas, há um comitê específico (*SPEC Virtualization Comitee*) trabalhando na definição de uma suíte de benchmark para a VIRTUALIZAÇÃO. Como produto gerado pelo

comitê foi criado recentemente o **SPECvirt_sc2010,** que pode ser encontrado no site SPEC.org. O SPECvirt_sc2010 é focado em benchmarks de servidores virtualizados e comentado posteriormente no Capítulo 9.

5.9. Limitações

As principais limitações e dificuldades apontadas para utilizar a VIRTUALIZAÇÃO são:

- **Aplicativos de carga excessiva**: aplicativos de carga excessiva, incluindo sistemas gerenciadores de bancos de dados, podem ser um fator limitante. Considerando que sempre existe uma perda de desempenho introduzida pelo HYPERVISOR, se uma aplicação ou um sistema gerenciador de banco de dados já demanda boa parte dos recursos do servidor, qual seria a razão para virtualizar este servidor?
- **Gerenciamento do licenciamento**: o gerenciamento do licenciamento pode ser um fator limitante. É necessário saber exatamente a regra para cada aplicação e isto é feito de maneira diferente pelos diversos fabricantes. Em uma determinada situação de carga, o licenciamento é válido, em outro, que utiliza uma outra configuração de hardware, pode não ser.
- **Falta de profissional especializado**: como a VIRTUALIZAÇÃO é relativamente nova, ainda existem poucos profissionais experientes que dominem a técnica e as opções comerciais disponíveis. Este aspecto deve ser considerado quando da escolha do software de VIRTUALIZAÇÃO.

5.10. Utilização

O software de VIRTUALIZAÇÃO pode ser utilizado de diversas maneiras para várias finalidades. Alguns dos principais usos são:

- Implementar a consolidação e a contenção dos servidores de produção. Conter a proliferação de servidores com a execução de aplicativos de software em máquinas virtuais implantadas em um menor número de servidores de nível corporativo, escaláveis e confiáveis. As taxas de consolidação chegam a dez ou mais máquinas virtuais por processador físico, aumentando drasticamente a utilização dos servidores e contendo sua proliferação.
- Propiciar a continuidade dos negócios por um custo reduzido. Proporcionar alta disponibilidade às aplicações críticas com soluções econômicas baseadas na VIRTUALIZAÇÃO. Pode-se implementar uma plataforma unificada de recuperação de desastres, permitindo que várias máquinas virtuais de produção sejam recuperadas no caso de uma falha de hardware, sem a necessidade de investir no caro mapeamento um-para-um entre o hardware de produção e o hardware de recuperação de dados.
- Simplificar o teste e o desenvolvimento de software. Consolidar ambientes distintos de desenvolvimento, teste e transferência, englobando diversos sistemas operacionais e aplicativos de vários níveis. Criar portais de autoatendimento para desenvolvedores, aumentando sua produtividade.

- Proteger e gerenciar desktops corporativos. Proteger os desktops corporativos de forças de trabalho geograficamente dispersas, fornecendo uma imagem padrão de desktop corporativo em máquina virtual. Ao mesmo tempo, fornecer ambientes padronizados de desktops corporativos, hospedados em máquinas virtuais acessadas via thin clients ou PCs.
- Simplificar o provisionamento de infraestrutura. Reduzir o tempo necessário para provisionar nova infraestrutura com recursos sofisticados de automação. Os dispositivos virtuais combinam a simplicidade da implementação de software com os benefícios dos dispositivos pré-configurados. Centralizar o controle e a responsabilidade pelos recursos de hardware, ao mesmo tempo atribuindo às unidades de negócios e aos proprietários de aplicações o total controle sobre a forma como os recursos são utilizados.

5.11. Referências Bibliográficas

Building the Virtualized Enterprise, VMware white paper, 2006.

ESG Research Report: The Impact of Server Virtualization on Storage, 2007.

Leveraging Microsoft Optimization to Create Your Dynamic IT Roadmap, Microsoft, 2009.

Magic Quadrant for x86 Server Virtualization Infrastructure. GARTNER RAS CORE RESEARCH NOTE, G00200526, may 2010.

Ruest, Daniel & Ruest, Nelson. Virtualization: A Beginner´s Guide. McGraw Hill, 2009.

Silberschatz, Abraham; Galvin, Peter & Gagne, Greg. Sistemas Operacionais: Conceitos e Aplicações. Elsevier, 2000.

The Business Value of Virtualization, Forrester Consulting, 2009.

Veras, Manoel. DATACENTER: Componente Central da Infraestrutura de TI, Brasport, 2009.

Virtualização: Eficiência sob medida, Executive Briefing, Computerworld, 2009.

Virtualization Overview, VMware white paper, 2006.

6. Técnicas

6.1. Introdução

Um sistema de computação pode ser dividido em quatro grandes componentes: hardware, sistema operacional (SO), aplicativos e usuários. Estes componentes, por sua vez, são divididos em subcamadas. Para melhor entender o princípio de funcionamento da VIRTUALIZAÇÃO e os tipos de máquinas virtuais possíveis de serem criadas, é preciso conhecer aspectos básicos de dois destes componentes: hardware e sistema operacional (SO).

O sistema de computação foi concebido utilizando uma hierarquia com diferentes níveis de abstração e com interfaces bem definidas entre esses níveis que o permite evoluir rapidamente.

O uso de níveis de abstração e interfaces, tanto para os componentes do SO como para os componentes de hardware, permitiu que cada componente fosse visto como um subsistema independente oferecendo serviços para os demais. Desta forma, os detalhes internos de implementação de cada um deles não precisam ser conhecidos: basta conhecer as interfaces e os serviços oferecidos. Ao mesmo tempo, ganha-se em simplificação e velocidade de atualização das plataformas.

6.2. Hardware

A arquitetura de computadores faz a descrição lógica e funcional dos componentes que formam o hardware de um sistema computacional e suas interações.

A arquitetura de computadores pode ser dividida em três grandes partes (CARISSIMI, 2009):

- **Conjunto de instruções de máquina *(Instruction Set Architecture – ISA):*** é a abstração do processador através de seu conjunto de instruções de máquina (*assembly*). Inclui, além das instruções, os modos de endereçamento possíveis e os registradores de máquina existentes.
- **Projeto do sistema**: envolve os componentes externos ao processador como barramentos, memória, controladores, arbitramento, subsistema de I/O e suas interconexões. Trata ainda dos mecanismos de suporte necessários a multiprocessadores, por exemplo.

- **Microarquitetura**: descrição de como são constituídas as unidades internas de um processador e como elas são interligadas para implementar o conjunto de instruções. Diz respeito a como um processador é organizado internamente.

Cada uma dessas partes possui um nível de abstração que fornece um conjunto de serviços e uma interface. O conjunto de instruções (ou ISA) é a interface entre o nível de abstração de hardware e o SO. Essa interface é composta por todos os códigos de máquina aceitos pelo processador, que correspondem, cada um, a uma instrução.

Tipicamente, um processador possui pelo menos dois modos de operação, não privilegiado e privilegiado. A interface ISA é dividida em dois subconjuntos para refletir cada um desses modos de operação.

- **User ISA**: o primeiro subconjunto é formado por todas as instruções de máquina que podem ser diretamente executadas por programas de usuário, as instruções de usuário ou user ISA. Os programas de usuário são executados em modo não privilegiado.
- **System ISA**: o segundo subconjunto é formado pelas instruções de sistema ou system ISA, que é constituído por aquelas instruções capazes de configurar o comportamento do próprio processador e de acessar componentes de hardware diretamente, como as instruções de I/O. Essas instruções são acessíveis unicamente em modo privilegiado e são executadas exclusivamente pelo núcleo (kernel) do sistema operacional.

6.3. Sistema Operacional (SO)

O SO é um alocador de recursos. Um sistema de computação possui muitos recursos que podem ser alocados para resolver um problema. O SO atua como gerente destes recursos e os aloca a programas e usuários específicos, conforme necessário, para execução das tarefas.

O programa de partida (*bootstrap program*) inicializa todos os aspectos do sistema e carrega o SO na memória do computador. O SO, em seguida, inicia a execução do primeiro processo, o "init", e espera que algum evento ocorra. A ocorrência de um evento é geralmente assinalada por uma interrupção de hardware ou software. Os SOs normalmente são baseados em interrupções e permitem que os recursos sejam compartilhados entre vários programas e processos diferentes e em tempo compartilhado.

"Sistemas Operacionais funcionam baseados em processo. Um processo é uma abstração que representa um programa em execução. Cada processo é um ambiente de execução isolado dos demais processos que executa sobre um processador lógico, isto é, um processador virtual, vinculado a si no momento da criação do processo. Cabe ao núcleo do sistema operacional, através de seu escalonador, alternar os diferentes processadores lógicos (virtuais) sobre um processador físico. A ilusão de paralelismo é criada pelo chaveamento rápido entre os processos." (CARISSIMI, 2009)

Para manter o controle sobre o sistema de computação, os desenvolvedores criaram o modo dual de execução privilegiado e não privilegiado. Este conceito dá suporte ao conceito de instruções privilegiadas que só podem ser executadas em modo privi-

legiado. Em certas situações um programa de usuário pode precisar realizar operações privilegiadas.

Um processo ou programa de usuário só pode empregar instruções não privilegiadas (*user ISA*), entretanto, em certas situações, como para realizar uma operação de I/O, os processos de usuário precisam executar instruções no modo privilegiado. Para permitir isso sem perder o controle, o sistema operacional disponibiliza aos programas de usuário um conjunto de chamadas de sistema (*system calls*).

As chamadas de sistema possibilitam que os programas de usuários acessem de forma indireta e controlada, através do sistema operacional, os recursos de hardware. Ao fazer isso, cabe ao sistema operacional realizar todos os procedimentos necessários para garantir que o processo de usuário não comprometa o correto funcionamento do sistema. As chamadas de sistema constituem, portanto, uma interface entre o processo de usuário e o sistema operacional, violando uma regra básica de um sistema computacional.

Um sistema operacional fornece um ambiente para a execução de programas. Fornece serviços para programas e para usuários dos programas. Os principais serviços são:

- Execução do Programa.
- Operações de I/O.
- Manipulação do Sistema de Arquivos.
- Comunicação.
- Detecção de Erro.

Um sistema operacional também possui um conjunto de bibliotecas de funções denominadas API – *Application Programming Interface*.

Uma API é definida para ser usada em conjunto com uma linguagem de programação de alto nível. Algumas APIs de bibliotecas apenas encapsulam chamadas de sistemas para apresentá-las de uma forma mais amigável ao programador, enquanto outras podem ser mais complexas e envolver várias chamadas de sistema para realizar uma determinada funcionalidade.

Para que os programas de usuário e as bibliotecas possam rodar sobre um sistema operacional e processador (plataforma), é preciso respeitar as chamadas de sistema disponibilizadas e ter um código binário compatível com o do processador. Isso é obtido através da compilação do programa e das bibliotecas para essa plataforma. O produto final é um código executável que contém os equivalentes binários das instruções de usuário (*user ISA*) e das chamadas de sistema (*system calls*).

Um sistema de computação é formado por camadas em uma hierarquia. O hardware é o nível mais baixo e o kernel (núcleo) executando no próximo nível utiliza instruções de hardware para criar um conjunto de camadas ao sistema para uso por camadas externas. Os processos, então, acima do kernel podem fazer chamadas ao sistema ou as instruções de hardware.

Usando técnicas de escalonamento de CPU e memória virtual, um sistema operacional pode criar a ilusão de que um processo tem seu próprio processador e memória. Este é o principio básico da VIRTUALIZAÇÃO.

A **Figura 6-1** ilustra os modelos de sistema de computação com máquina não virtual e virtual.

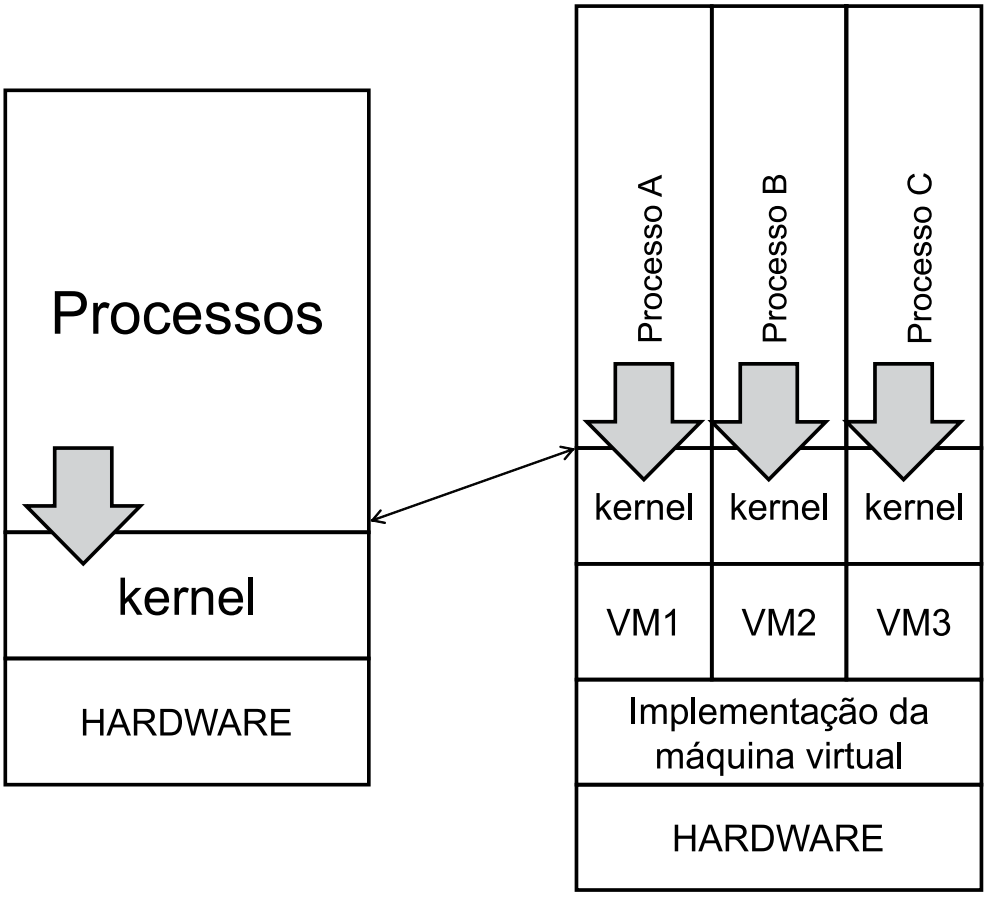


Figura 6-1 Modelos de Sistema (a) Não Virtual e (b) Virtual

6.4. Categorias de Virtualização

Uma forma de entender a VIRTUALIZAÇÃO é entender que ela consiste em estender ou substituir um recurso ou uma interface existente por outro, de modo a imitar um comportamento. Isso é feito através de uma camada de software responsável por transformar ações de um sistema A em ações equivalentes em um sistema B. Dependendo de como e onde essa transformação é feita, é possível classificar os softwares de VIRTUALIZAÇÃO em três grandes categorias:

- **Nível de hardware:** a camada de VIRTUALIZAÇÃO é posta diretamente sobre a máquina física e a apresenta às camadas superiores como um hardware abstrato similar ao original. Corresponde à definição original de máquina virtual dos anos 60.

- **Nível de sistema operacional:** a camada de VIRTUALIZAÇÃO é um mecanismo que permite a criação de partições lógicas em uma plataforma de maneira que cada partição seja vista como uma máquina isolada, mas que compartilha o mesmo sistema operacional. Nesse caso, a camada de VIRTUALIZAÇÃO se insere entre o sistema operacional e as aplicações e é a categoria que nos interessa.
- **Nível de linguagens de programação:** a camada de VIRTUALIZAÇÃO é um programa de aplicação do sistema operacional. O objetivo é definir uma máquina abstrata sobre a qual executa uma aplicação desenvolvida em uma linguagem de programação de alto nível específica.

6.5. Máquina Virtual de Sistema ou HYPERVISORS

Neste caso a máquina virtual oferece um ambiente de execução completo onde podem coexistir um sistema operacional e vários processos, possivelmente de diferentes usuários. Dessa forma, uma única plataforma de hardware pode executar múltiplos sistemas operacionais hóspedes, um em cada máquina virtual, simultaneamente.

O HYPERVISOR é a plataforma de máquina virtual. Suas principais funções consistem em agendamento, gerência da memória e manutenção do estado da máquina virtual. Também permite criar partições para as máquinas virtuais, mantendo o isolamento entre as partições.

O desempenho e a escalabilidade do HYPERVISOR definem a qualidade da VIRTUALIZAÇÃO. Segurança sobre os recursos virtualizados e agilidade de reconfigurar recursos computacionais sem interromper as operações do servidor de máquinas virtuais são algumas das características necessárias ao HYPERVISOR.

Os HYPERVISORS para efeito de simplificação podem ser classificados em dois tipos:

6.5.1. TIPO UM – Baremetal

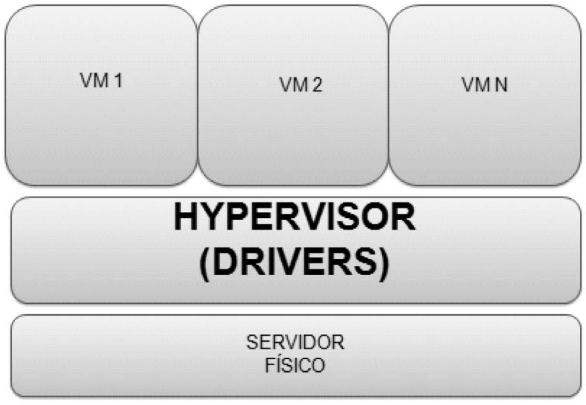
HYPERVISOR que roda diretamente no hardware do servidor, também conhecido como BAREMETAL. Controla o hardware e o acesso do SO convidado (guest). O papel do HYPERVISOR nativo é compartilhar os recursos de hardware entre as máquinas virtuais de forma que cada uma delas imagina ter recursos exclusivos. São aqueles que rodam diretamente sobre o hardware de uma máquina real e as máquinas virtuais são postas sobre ele. A função básica de um HYPERVISOR nativo é compartilhar os recursos de hardware (processador, memória, meios de armazenamento e dispositivos de I/O) entre as diferentes máquinas virtuais de forma que cada uma delas tenha a ilusão de que esses recursos são privativos a ela.

Esse tipo de HYPERVISOR corresponde ao originalmente implementado nos sistemas IBM, no início da década de 70. Exemplos: VMware ESX Server, Microsoft Hyper-V e Citrix Xen Server. Uma variação do TIPO UM é o embeeding HYPERVISOR, que é instalado no firmware (como o VMware ESXi). Este HYPERVISOR é de pequeno tamanho e tem um impacto mínimo nos recursos e no desempenho do servidor físico.

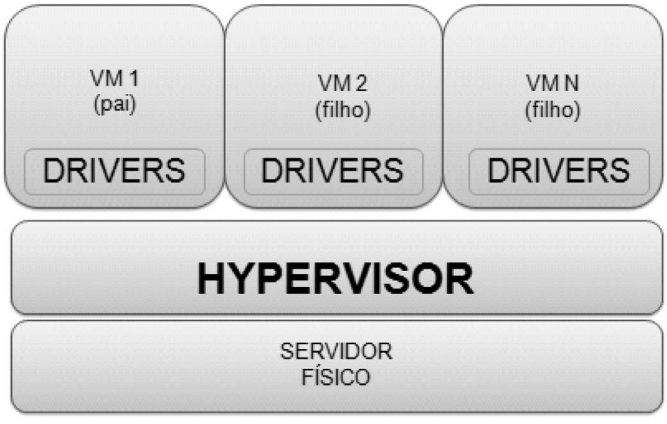
Existem dois tipos de HYPERVISORS TIPO 1: HYPERVISOR Monolítico e HYPERVISOR microkernelizado, conforme ilustra a **Figura 6-2**.

O HYPERVISOR monolítico, **Figura 6-2 (a),** precisa de uma grande quantidade de código entre os recursos de hardware e as VMs, porque este tipo de HYPERVISOR emula todo o hardware para as VMs. Nesta opção os *drivers* estão no próprio HYPERVISOR.

O HYPERVISOR microkernelizado, **Figura 6-2 (b),** utiliza *drivers* na própria máquina virtual e a única camada entre o SO convidado e o hardware é o HYPERVISOR. Este HYPERVISOR não utiliza drivers de terceiros. Os drivers estão na própria máquina virtual. Este tipo de HYPERVISOR tem mais segurança em sua arquitetura, devido a sua superfície de ataque mínima.



(a)



(b)

Figura 6-2 Arquitetura Monolítica (a) versus Microkernelizada (b)

6.5.2. TIPO DOIS

É uma aplicação que fornece um ambiente de execução para outras aplicações. Roda sobre um sistema operacional nativo como se fosse um processo deste. A camada de VIRTUALIZAÇÃO é composta por um sistema operacional convidado possivelmente diferente do sistema operacional nativo, e por um hardware virtual criado sobre os recursos de hardware oferecidos através do SO nativo. Um exemplo é a máquina virtual Java (JVM).

Existem algumas soluções de VIRTUALIZAÇÃO que utilizam um HYPERVISOR híbrido com características dos dois tipos mencionados.

A implementação das máquinas virtuais não é uma tarefa tão simples quanto parece. Além da preocupação direta com desempenho, existe o problema de como os recursos físicos da máquina são compartilhados entre o sistema nativo e o sistema convidado sem que um interfira em outro, a começar pelo próprio processador. O problema fundamental consiste no que fazer quando o sistema convidado executa instruções privilegiadas (System ISA), já que essas, por uma questão de proteção do sistema, são exclusivas do sistema nativo. A ação a ser tomada depende de recursos oferecidos pela arquitetura do processador (hardware). Como teorema aceito em computação tem-se que qualquer instrução que possa afetar o comportamento do sistema deve ser monitorada e tratada adequadamente.

Para melhor entender o problema do privilégio da execução das instruções é interessante entender a arquitetura x86. A arquitetura x86 provê quatro modos de operação para o processador, identificados de 0 a 3, denominados de anéis de proteção (Rings) ou CPL (Current Privilege Level). Nos sistemas operacionais convencionais (Microsoft Windows e UNIX), para esse tipo de arquitetura apenas dois modos são usados: o Ring 0, que detém os maiores privilégios de execução, é usado pelo sistema operacional, e o Ring 3, de menor privilégio, é empregado pelos processos de usuário. Se um processo de usuário tentar executar uma instrução privilegiada ocorrerá uma exceção (trap) que deverá ser tratada adequadamente. Entretanto, a arquitetura x86 possui dezessete instruções sensíveis que não são privilegiadas em termos de arquitetura, mas que são privilegiadas na prática e que podem acessar o processador diretamente sem gerar exceções (traps). Esta era a grande dificuldade com a VIRTUALIZAÇÃO na plataforma x86.

A VMware resolveu o desafio em 1998 desenvolvendo técnicas de translação binária que possibilitaram o VMM rodar no Ring 0 para isolação e desempenho, enquanto moveu o SO para um nível de usuário com mais privilégios de que as aplicações de usuário em Ring 3, mas com menos privilégios do que o VMM no Ring 0.

Na prática, fazer a VIRTUALIZAÇÃO funcionar na arquitetura x86 é alterar os níveis de privilégios padrões de execução da arquitetura x86. A **Figura 6-3** ilustra a arquitetura típica de um processador.

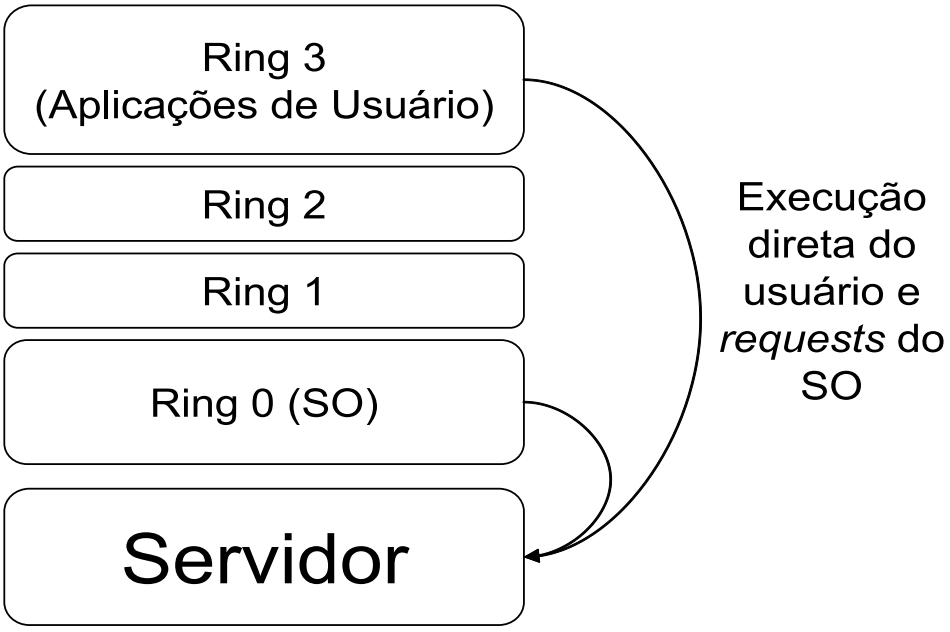


Figura 6-3 Hierarquia na Arquitetura de processador x86

A VIRTUALIZAÇÃO pode ser classificada em três tipos na arquitetura x86:

- VIRTUALIZAÇÃO TOTAL baseada em translação binária.
- PARAVIRTUALIZAÇÃO ou VIRTUALIZAÇÃO assistida pelo SO.
- VIRTUALIZAÇÃO assistida pelo HARDWARE.

6.6. VIRTUALIZAÇÃO Total

A VIRTUALIZAÇÃO TOTAL realiza a completa abstração do sistema físico e cria um sistema virtual completo. Não é necessário fazer qualquer modificação no SO ou na aplicação que está rodando nesta modalidade. Este tipo de VIRTUALIZAÇÃO facilita a migração de máquinas virtuais entre servidores físicos, pois existe total independência das aplicações e dos recursos físicos do servidor. Também a segurança é facilitada pela isolação entre as máquinas virtuais. O desempenho neste caso pode ser prejudicado, pois o HYPERVISOR controla todo processo e toda a chamada ao hardware é feita sob a sua supervisão. Também a implementação de uma máquina virtual que imite cada dispositivo de hardware é uma tarefa complexa, pois isto é feito baseado em hardwares genéricos, o que influi no desempenho.

A VIRTUALIZAÇÃO TOTAL é uma combinação de técnicas de translação binária e execução direta. O sistema operacional convidado é completamente abstraído e isolado do hardware. A VIRTUALIZAÇÃO TOTAL não requer qualquer modificação no hardware ou no sistema operacional. A **Figura 6-4** ilustra a arquitetura de privilégios no x86 para a VIRTUALIZAÇÃO TOTAL. Observe que o sistema operacional convidado roda no Ring 1.

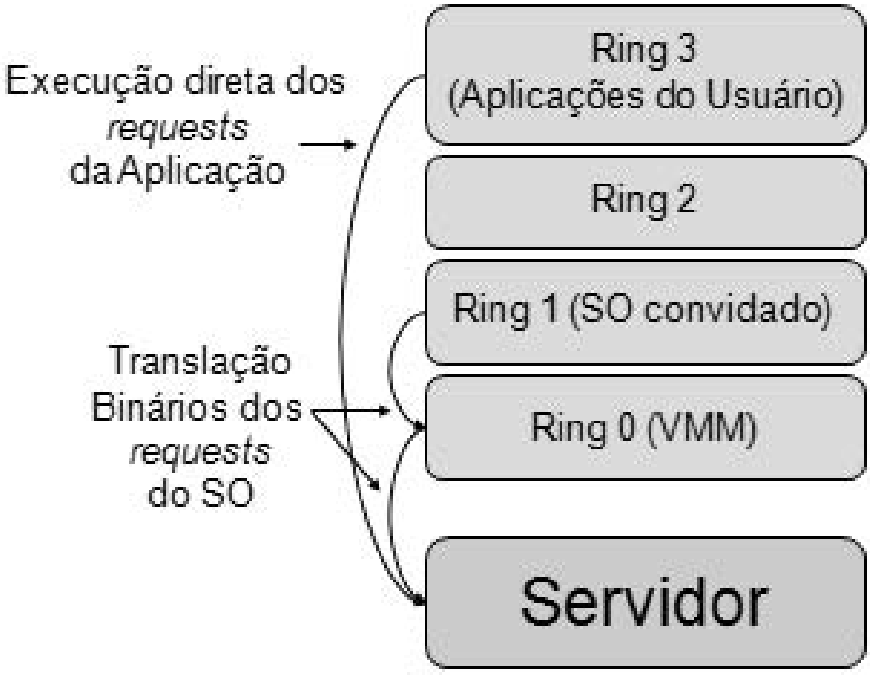


Figura 6-4 VIRTUALIZAÇÃO TOTAL no x86

Existem alguns inconvenientes para a VIRTUALIZAÇÃO TOTAL:

- As instruções, por não serem modificadas, precisam ser testadas pelo HYPERVISOR para saber se são sensíveis ou não. As instruções sensíveis devem ser interceptadas e emuladas no sistema nativo, para evitar que a máquina virtual altere o comportamento do sistema.
- A máquina virtual possui suporte de conjunto genérico de dispositivos devido à diversidade de dispositivos existentes. Pode-se ter subutilização de recursos quando comparado ao hardware real. Os dispositivos são genéricos e, portanto, perdem desempenho.
- Existem alguns problemas técnicos relativos à implementação da gerência de memória.

A VIRTUALIZAÇÃO TOTAL consiste em prover uma réplica (virtual) do hardware subjacente de tal forma que o sistema operacional e as aplicações possam executar como se estivessem diretamente sobre o hardware original. A grande vantagem é que o sistema operacional convidado não precisa ser modificado para rodar sobre o monitor de máquina virtual (VMM) ou HYPERVISOR. A **Figura 6-5** ilustra a arquitetura na VIRTUALIZAÇÃO TOTAL.

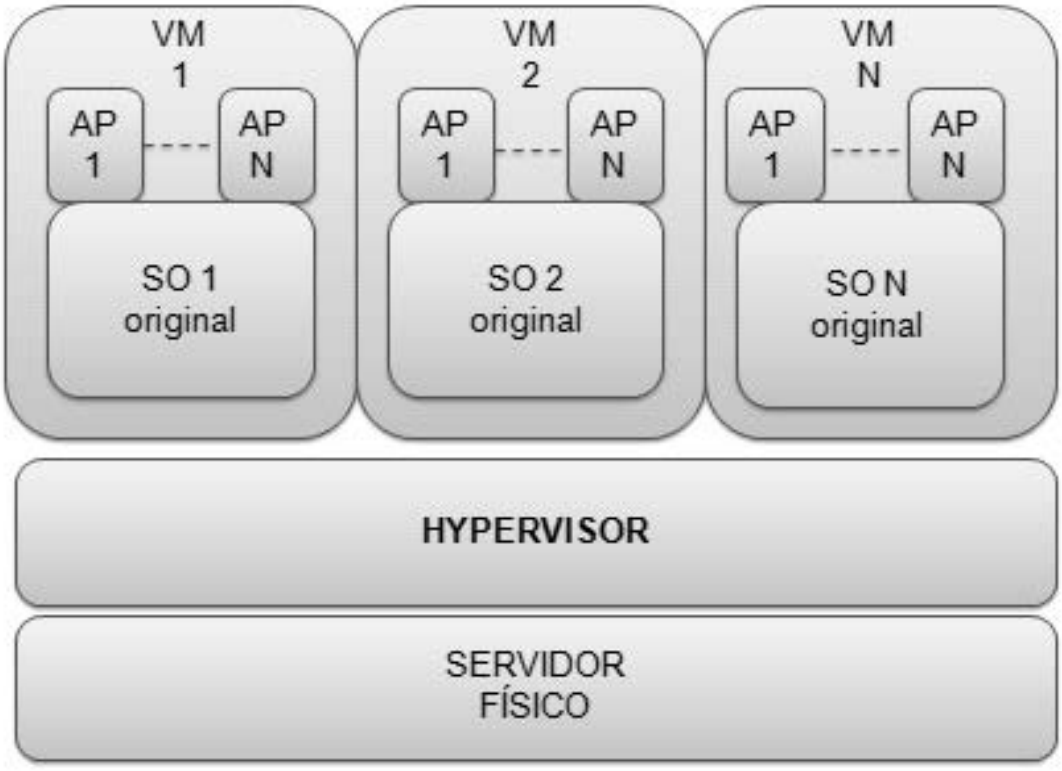


Figura 6-5 Arquitetura para a VIRTUALIZAÇÃO TOTAL

6.7. PARAVIRTUALIZAÇÃO

A PARAVIRTUALIZAÇÃO surgiu como uma forma de contornar as desvantagens de uso da VIRTUALIZAÇÃO completa, no que diz respeito ao processamento. A máquina virtual enxerga uma abstração do hardware que não é idêntico ao hardware físico. Os dispositivos de hardware são acessados por drivers de dispositivo do próprio HYPERVISOR, o que é interessante, pois otimiza o desempenho. O problema é que a PARAVIRTUALIZAÇÃO requer modificação do sistema operacional convidado.

O Xen Open Source é um exemplo de sistema baseado em PARAVIRTUALIZAÇÃO que virtualiza o processador e a memória, usando o núcleo modificado do Linux e virtualizando o I/O com drivers de dispositivos customizados.

A PARAVIRTUALIZAÇÃO é uma alternativa para contornar os problemas de desempenho da VIRTUALIZAÇÃO TOTAL. No caso da PARAVIRTUALIZAÇÃO, o sistema convidado é alterado para chamar a máquina virtual sempre que for executar uma instrução sensível. As instruções de usuário não precisam ser alteradas e podem ser executadas diretamente sobre o processador nativo. A **Figura 6-6** ilustra a PARAVIRTUALIZAÇÃO.

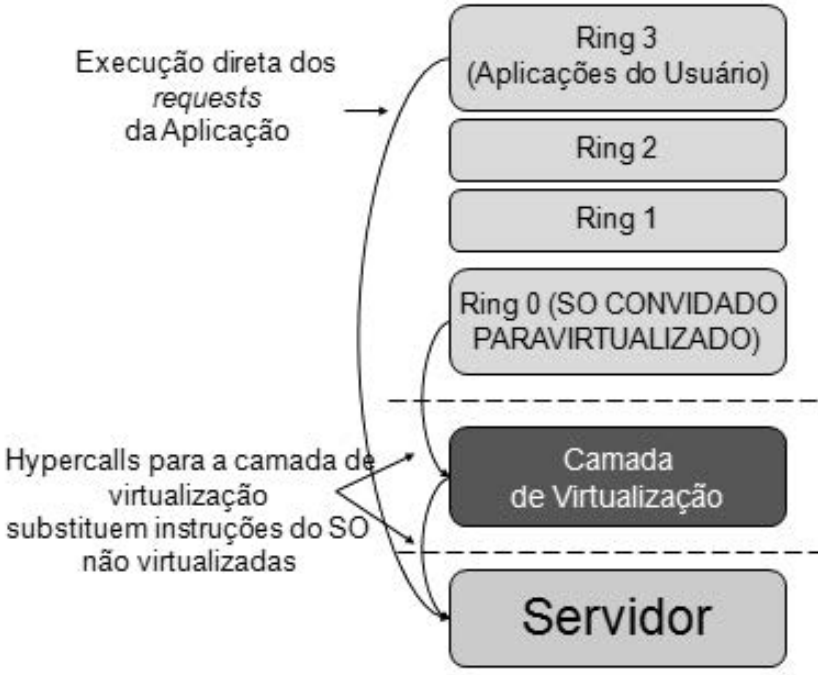


Figura 6-6 PARAVIRTUALIZAÇÃO no x86

A arquitetura da PARAVIRTUALIZAÇÃO é mostrada na **Figura 6-7**. Observe que os sistemas operacionais convidados são modificados. Sem dúvida, este é um aspecto da PARAVIRTUALIZAÇÃO que dificulta a sua adoção. Os fornecedores de sistema operacional deveriam oferecer duas versões dos seus sistemas, uma delas adequada à PARAVIRTUALIZAÇÃO.

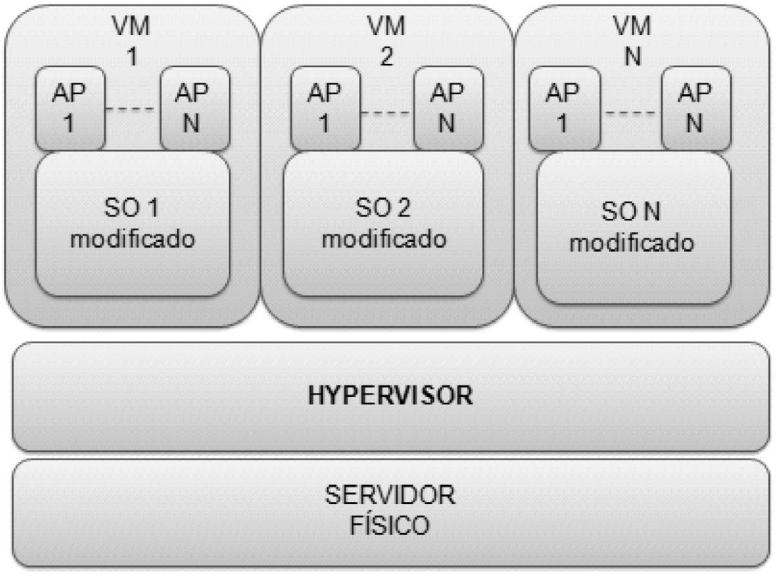


Figura 6-7 Arquitetura para a PARAVIRTUALIZAÇÃO

6.8. VIRTUALIZAÇÃO Assistida por Hardware

Os fabricantes Intel e AMD investiram em extensões na arquitetura x86 para suportar a VIRTUALIZAÇÃO e melhorar o desempenho da solução como um todo.

Este movimento da AMD e da Intel praticamente eliminou as vantagens de desempenho dos sistemas baseados em PARAVIRTUALIZAÇÃO que tinham o *penalty* de ter que modificar o SO para funcionar. Em relação aos sistemas do tipo VIRTUALIZAÇÃO TOTAL utilizado pela VMware, a primeira geração deste tipo de VIRTUALIZAÇÃO perdia em desempenho para a VIRTUALIZAÇÃO TOTAL baseada em translação binária, mas os fabricantes continuam melhorando a performance do processador x86 em ambientes virtualizados. Em função da rigidez de programação, só novos sistemas x64 permitem que sistemas originalmente desenvolvidos para VIRTUALIZAÇÃO TOTAL façam também a VIRTUALIZAÇÃO assistida pelo hardware e assim melhorem o desempenho da VIRTUALIZAÇÃO.

No nível do processador tanto Intel (Intel VT) como AMD (AMD -V) fizeram um esforço para alterar os modos de operação do processador, os anéis (Rings) de proteção identificados de 0 a 3. O HYPERVISOR passou a rodar em um anel abaixo do Ring 0, criado especificamente para melhorar o desempenho de servidores x86 virtualizados. Ou seja, o HYPERVISOR passou a ter total prioridade sobre o sistema operacional. A **Figura 6-8** ilustra a VIRTUALIZAÇÃO assistida pelo hardware.

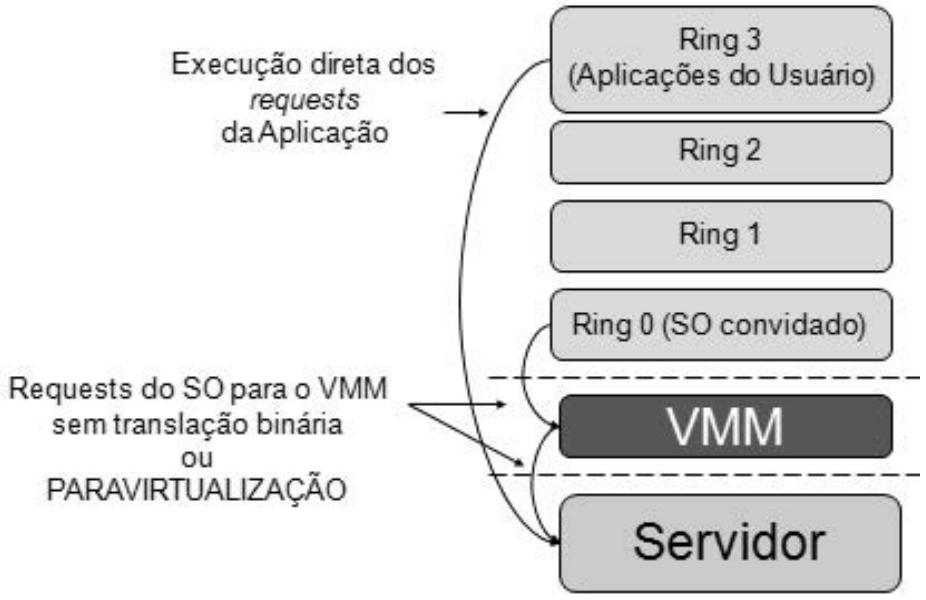


Figura 6-8 Hardware Assist no x86

6.9. Comparação entre as Técnicas de VIRTUALIZAÇÃO

A **Tabela 6-1** faz um resumo do que foi tratado. Observe que os principais softwares de mercado, na verdade, utilizam uma combinação de técnicas e tiram o melhor proveito de cada uma delas.

Tabela 6-1 Sumário com comparação sobre as técnicas de VIRTUALIZAÇÃO

	VIRTUALIZAÇÃO TOTAL	VIRTUALIZAÇÃO ASSISTIDA POR HARDWARE	PARA VIRTUALIZAÇÃO
Técnica	Translação Binária e Execução Direta	Saída para modo raiz nas instruções privilegiadas	*Hypercalls*
Modificação do SO Convidado/ Compatibilidade	SO convidado não modificado/excelente	SO convidado não modificado/excelente	SO convidado modificado/baixa compatibilidade
Desempenho	Bom	Considerável	Melhor em certos casos
Usado por	VMware, Microsoft, Parallels	VMware, Microsoft, Parallels, Xen	VMware, Xen
Independência entre SO convidado e VMM	Sim	Sim	Não

6.10. Referências Bibliográficas

Carissimi, Alexandre. Virtualização: Princípios Básicos e Aplicações, UFRGS, 2009.

Rosemblum, M. The Reincarnation of Virtual Machines. Queue Focus – ACM Press, 2004.

Silberschatz, Abraham; Galvin, Peter & Gagne, Greg. Sistemas Operacionais: Conceitos e Aplicações. Elsevier, 2000.

Understanding Full Virtualization, Paravirtualization and Hardware Assist. VMware white paper, 2007.

7. Arquitetura e Infraestrutura

7.1. Introdução

O software de VIRTUALIZAÇÃO exige uma arquitetura e uma infraestrutura de apoio que pode ser muito simples em pequenas instalações, mas também pode ser muito complexa em grandes instalações.

Os principais softwares do mercado como o VMware VSphere e o Microsoft WS08 Hyper-V demandam uma infraestrutura que deve permitir obter os níveis de serviço adequados para as aplicações em termos de desempenho, disponibilidade e segurança. Como regra geral, o software de VIRTUALIZAÇÃO deve otimizar a demanda de processamento, de armazenamento e de I/O, distribuindo a carga de trabalho para otimizar o uso dos recursos.

A infraestrutura de TI existente deve ser avaliada em novos projetos de VIRTUALIZAÇÃO. Alguns componentes da infraestrutura como servidores e unidades de armazenamento já utilizados pela organização podem não ser aproveitados em um projeto de VIRTUALIZAÇÃO. Também quando da migração para o novo ambiente deve-se considerar a existência de ferramentas que automatizam a migração das máquinas físicas (P) para máquinas virtuais (V), a chamada operação P2V. Estas ferramentas são disponibilizadas pelos principais fabricantes de software. Também a definição de métricas de desempenho é uma questão relevante para monitoração da qualidade da infraestrutura de TI, antes e depois da VIRTUALIZAÇÃO.

Um projeto de VIRTUALIZAÇÃO traz implicações para os outros componentes da TI e para a própria arquitetura utilizada no DATACENTER. A VIRTUALIZAÇÃO altera alguns pressupostos comuns em um projeto convencional e introduz modificações profundas na maneira de pensar o projeto do DATACENTER.

A **Figura 7-1** sinaliza os efeitos da VIRTUALIZAÇÃO nos principais componentes do DATACENTER.

Figura 7-1 Impacto da VIRTUALIZAÇÃO

Também a VIRTUALIZAÇÃO traz impactos para o projeto da engenharia do DATACENTER, pois altera o mapa da densidade de energia e a consequente refrigeração do DATACENTER. As cargas de trabalho passam a se movimentar dinamicamente com a VIRTUALIZAÇÃO e alteram o perfil energético do DATACENTER.

7.2. Arquitetura Física do DATACENTER

Normalmente o DATACENTER, do ponto de vista da TI, é construído com base em uma arquitetura hierárquica que segue o modelo utilizado em projetos de redes, só alterando a camada de distribuição pela camada de agregação.

A **Figura 7-2** ilustra a arquitetura em camadas do DATACENTER.

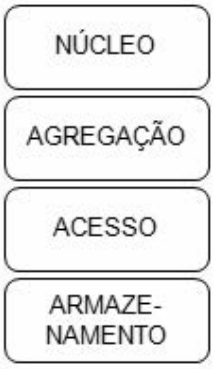


Figura 7-2 Camadas no DATACENTER

O DATACENTER físico pode ser representado pelo diagrama mostrado na **Figura 7-3**. A Cisco afirma que esta arquitetura baseada em camadas tem sido testada e validada por vários anos em grandes DATACENTERS espalhados pelo mundo.

As camadas de rede do modelo hieráquico são:

- **NÚCLEO**: formada por switches de alta velocidade para os fluxos que entram e saem do DATACENTER.
- **AGREGAÇÃO**: formada por switches com função de integração dos principais serviços de rede incluindo balanceamento de carga, detecção de intrusão, firewalls, SSL offload, etc.
- **ACESSO**: formada por switches onde os servidores são conectados à rede e onde as políticas de rede (ACLs, QoS, VLANs) são implementadas.

A camada de acesso pode ser implementada com grandes switches modulares localizados no fim de cada coluna localizada no DATACENTER (Modelo *End-of-Row* – EoR) ou com switches com configuração fixa no topo dos RACKS (Modelo *Top-of-Rack* – ToR), que propiciam conectividade para um ou alguns RACKS adjacentes e possuem link para a camada de agregação.

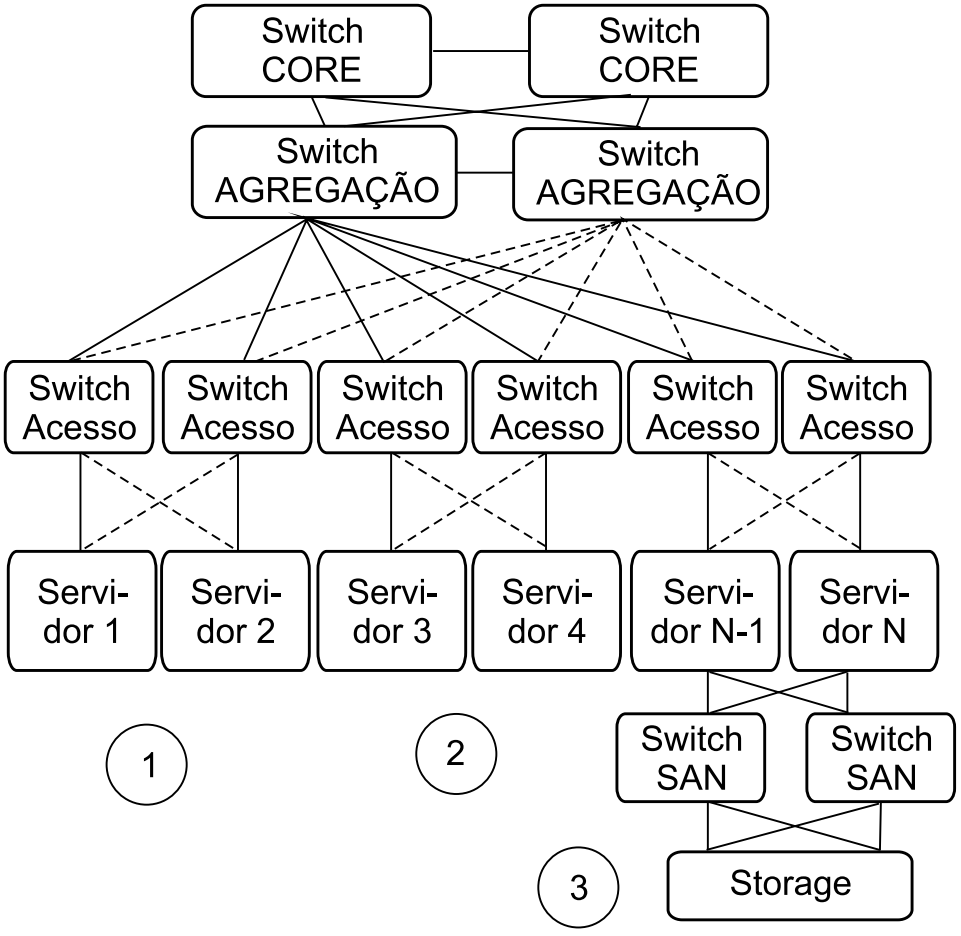


Figura 7-3 Arquitetura Convencional do DATACENTER

O avanço da tecnologia Ethernet tem possibilitado a unificação das redes dentro do DATACENTER. Uma tendência é a de que toda a comunicação interna seja baseada em dispositivos que se conectam via Ethernet. A previsão de aumento de velocidade deste padrão de 1GbE para 10GbE, 40GbE e 100GbE nos próximos dez anos reforça a sua utilização para conectar servidores.[9]

Os principais recursos disponibilizados por uma arquitetura do DATACENTER físico são:

- Recursos de processamento e memória, incluindo servidores e clusters.
- Recursos de armazenamento.
- Recursos de conectividade.

Dois efeitos recentes alteram o perfil de arquitetura e a forma de disponibilizar recursos até então utilizados no DATACENTER.

- A introdução da arquitetura baseada em blades produz mudanças profundas no DATACENTER, alterando o perfil de utilização do espaço físico, o gerenciamento, a comunicação e a refrigeração.
- A VIRTUALIZAÇÃO invalida ou pelo menos altera a arquitetura convencional mostrada anteriormente, pois altera a relação de uma imagem de sistema operacional para um servidor (1:1) para várias imagens para um servidor (n:1).

7.3. Arquitetura Virtual do DATACENTER

A VIRTUALIZAÇÃO possibilita otimizar o uso da infraestrutura de TI, incluindo servidores, storage e os dispositivos de rede. Ela agrega os vários recursos e apresenta um simples e uniforme conjunto de elementos em um ambiente virtual. O DATACENTER virtual pode então ser provisionado para o negócio com o conceito de infraestrutura compartilhada virtualmente.

Os principais recursos da arquitetura do DATACENTER virtual são:

- Recursos de processamento e memória, incluindo servidores, clusters e RPs (*resource pools* ou pool de recursos).
- Recursos de armazenamento e *datastores*.
- Recursos de conectividade.

Os servidores virtuais representam os recursos virtuais de processamento e memória de um servidor físico rodando o software de VIRTUALIZAÇÃO. O servidor é onde estão alojadas as máquinas virtuais (VMs).

Os clusters virtuais são o conjunto de servidores que possibilitam agregar dinamicamente recursos de processamento e memória. Estes clusters atuam de forma coordenada e permitem obter funcionalidades que não são obtidas com um servidor isolado.

9 X86 server forecast by ethernet connection type no IEEE 802.3 HSSG, Abril 2007, Interim Meeting.

Os recursos de processamento e memória incluídos em servidores e clusters podem ser particionados em uma hieraquia de pool de recursos.

O storage é onde as aplicações, os dados e as informações de configuração ficam armazenados. Os *datastores* são representações virtuais de combinações de recursos físicos de storage.

As máquinas virtuais (VMs) são associadas a um servidor virtual particular, um cluster, um pool de recursos e um *datastore*. Provisionar VMs é mais simples do que provisionar servidores físicos. Os recursos são provisionados para as VMs baseados nas políticas definidas pelo administrador de sistemas, obtendo uma grande flexibilidade no novo ambiente. O pool de recursos pode ser reservado para uma VM específica, por exemplo. O pool de recursos permite alocar recursos específicos para as aplicações e funciona dentro de um cluster das máquinas virtuais.

A **Figura 7-4** ilustra a relação entre o DATACENTER físico e o DATACENTER virtual. Os recursos de hardware passam a fazer parte de verdadeiros pools de recursos virtuais utilizados de acordo com os níveis de serviço requeridos pelas aplicações.

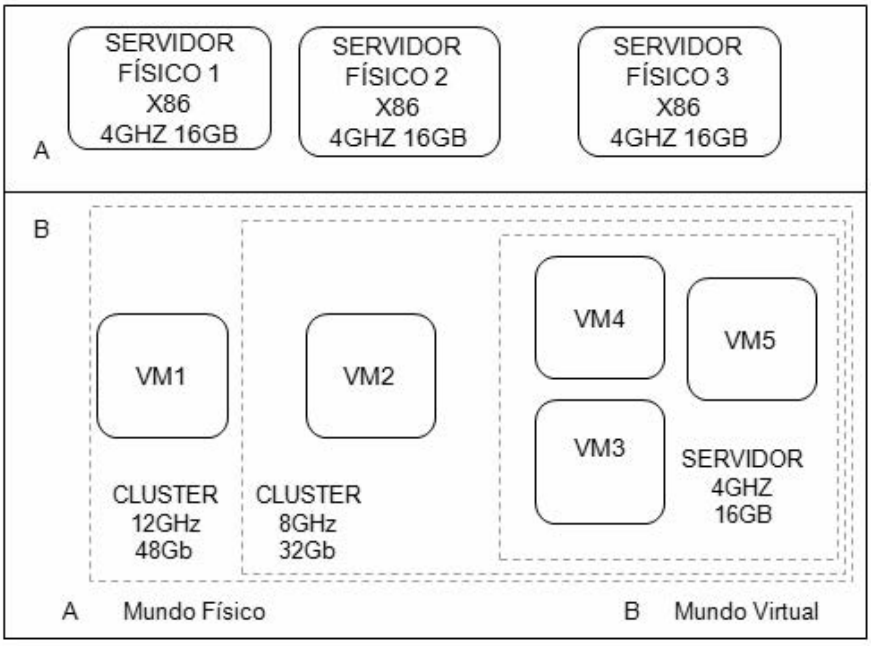
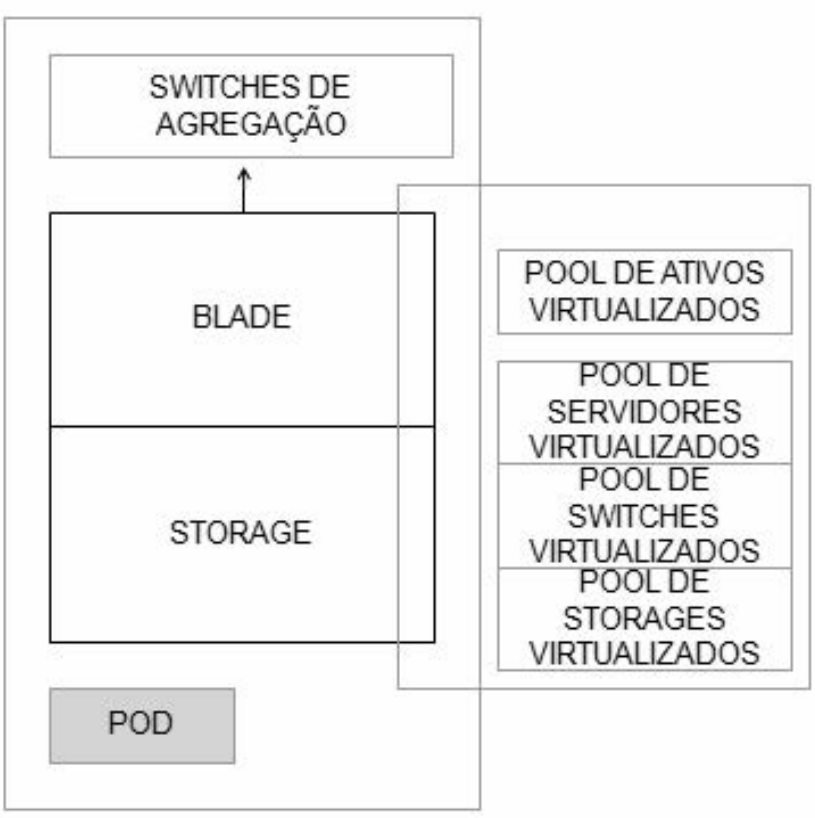
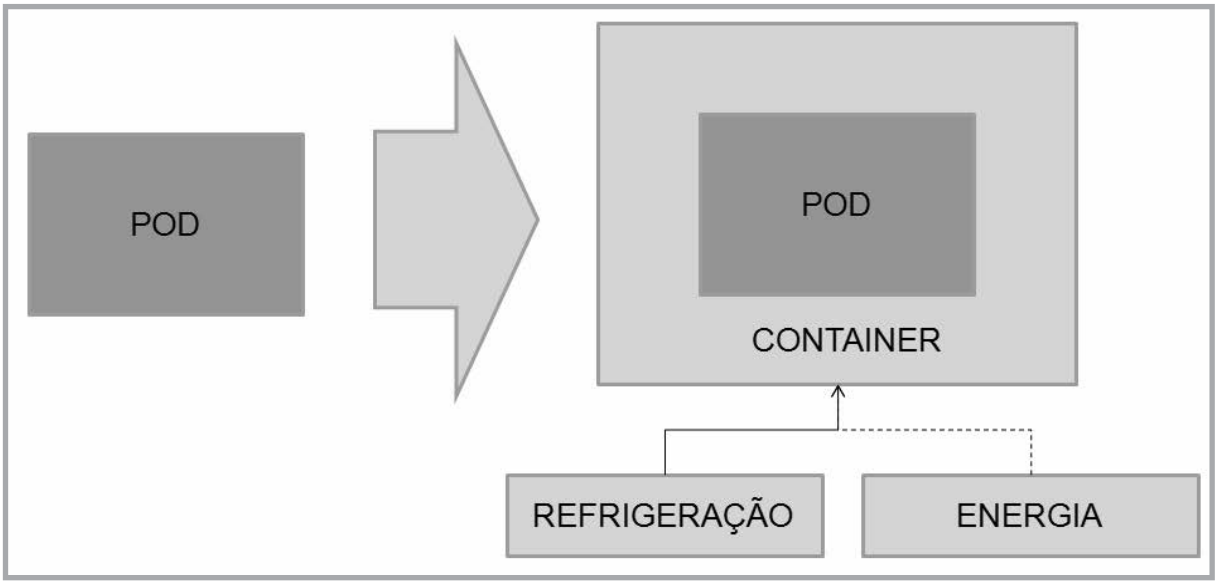


Figura 7-4 DATACENTER Físico e Virtual

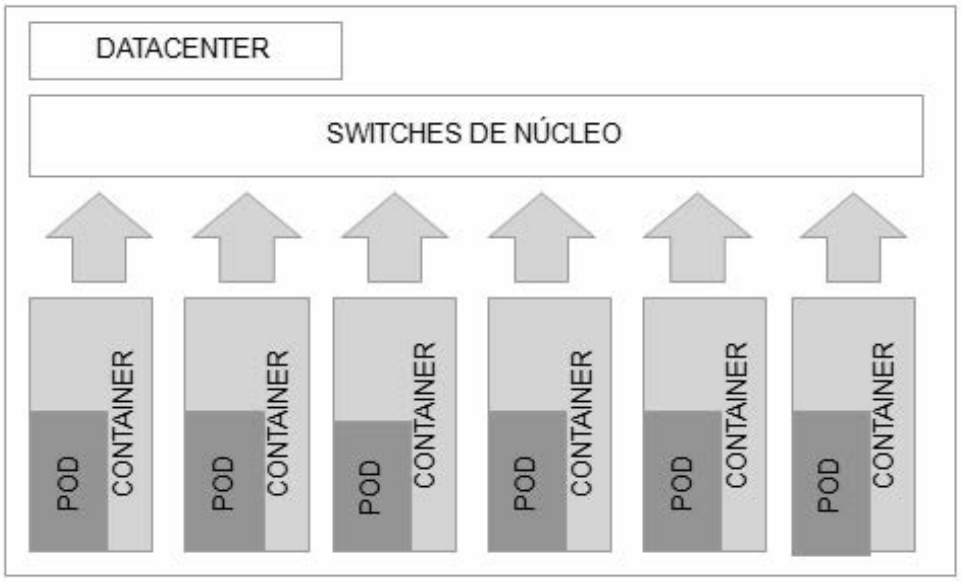
O DATACENTER virtualizado é um conjunto de tecnologias (*computing pods*) que funcionam como grandes conjuntos de recursos, conforme ilustra a **Figura 7-5 (a)**. Cada pool de recursos é representado por um conjunto de blades, storage e rede. Para cada um destes conjuntos existe um sistema de energia e refrigeração projetado para otimizar o uso dos recursos. Para cada conjunto de tecnologias existe uma plataforma simplificada de gerenciamento e provisionamneto. Os CONTAINERS podem ser os elementos físicos onde os PODs estão instalados, conforme ilustra a **Figura 7-5 (b).** Um conjunto de CONTAINERS forma o DATACENTER, conforme ilustra a **Figura 7-5 (c)**.



(a) POD



(b) CONTAINER e POD



(c) DATACENTER MODULAR

Figura 7-5 CONTAINERS, PODS e DATACENTER

Na arquitetura convencional os servidores rodam sistemas operacionais exclusivos que utilizam dispositivos de I/O físicos ligados ao servidor. A VIRTUALIZAÇÃO modifica esta arquitetura, na medida em que invalida a ideia central de um servidor físico rodando um único sistema operacional.

A VIRTUALIZAÇÃO também invalida a ideia de que a relação entre o sistema operacional e a rede é estática. Como a VIRTUALIZAÇÃO abstrai o hardware do software, ela permite que a imagem do SO, agora uma máquina virtual, possa ser movimentada entre servidores ou mesmo entre DATACENTERS. Do ponto de vista da rede, significa que os serviços implementados na camada de agregação precisam ser modificados para suportar a mobilidade das máquinas virtuais. Mesmo na camada de acesso, a mobilidade das máquinas virtuais pode trazer algumas novas necessidades que vão ao encontro das melhores práticas de redes estáticas.

7.4. Blades, Servidores e VIRTUALIZAÇÃO

7.4.1. Introdução

Os servidores utilizados nos DATACENTERS variam desde mainframes, servidores cujos processadores são baseados em instruções RISC (*reduced instruction set computer*) até servidores x86, cujos processadores são baseados em instruções CISC (*complex instruction set computer*).

Os servidores cujos processadores são baseados em instruções do tipo CISC contêm um grande número de instruções. Os modelos baseados em instruções RISC deixam para o sistema operacional boa parte da execução das instruções e, portanto, possuem número reduzido de instruções. Os modelos CISC, especificamente os modelos

que utilizam a arquitetura x86, são considerados um padrão da indústria e utilizados em larga escala, sendo o foco da VIRTUALIZAÇÃO.

Os servidores x86 vêm evoluindo rapidamente em termos de confiabilidade e poder de processamento, de modo que o seu uso hoje já acontece em grande escala nos mais diversos ambientes computacionais. A arquitetura desses servidores foi modificada recentemente com a introdução dos servidores de lâminas, os blades.

A arquitetura de servidores baseados no processador x86 evoluiu ao longo do tempo. Na medida em que novos processadores com maior poder de processamento são lançados, memória e barramento local (I/O) naturalmente acompanham esta evolução para garantir um sistema balanceado e sem gargalos.

Com a VIRTUALIZAÇÃO, os fabricantes Intel e AMD fizeram um esforço enorme de rapidamente compensar no hardware a perda de desempenho (overhead) ocasionada pela virtualização. O resultado disto é que novos processadores permitem a otimização do desempenho mesmo com a camada de virtualização instalada.

Um aspecto importante em um projeto de VIRTUALIZAÇÃO é que o hardware dos fabricantes precisa estar homologado para as soluções de virtualização encontradas no mercado.

O software de VIRTUALIZAÇÃO é intensivo em CPU, portanto:

- Deve-se sempre considerar a aquisição de processadores de maior desempenho, considerando a maior opção multicore.
- Os requisitos de CPUs das VMs são distribuídos através dos núcleos (cores).
- Cada VM pode suportar determinado número de virtual CPUs, dependendo do software utilizado.

O software de VIRTUALIZAÇÃO é intensivo em memória, portanto:

- Deve-se considerar para cálculo do tamanho da memória o número de VMs, a quantidade de memória por VM e a capacidade adicional para migração.
- A memória consumida pelo HYPERVISOR varia com o número de VMs e com a memória alocada para cada VM.
- Os fabricantes de servidores como Dell, HP, IBM possuem diversas configurações homologadas que suportam os softwares de VIRTUALIZAÇÃO.
- Para o software de gerenciamento deve-se adquirir uma configuração de hardware suportada, normalmente o sistema operacional Windows.

As configurações de servidores a serem utilizadas para a VIRTUALIZAÇÃO variam entre as modalidades *scale-out* e *scale-up*. Ou seja, a VIRTUALIZAÇÃO pode ser adotada em pequenos servidores que crescem horizontalmente (aumento do número de servidores) ou em grandes servidores com crescimento vertical (aumento do número de processadores). Na decisão por qual modalidade adotar, deve-se verificar de que forma

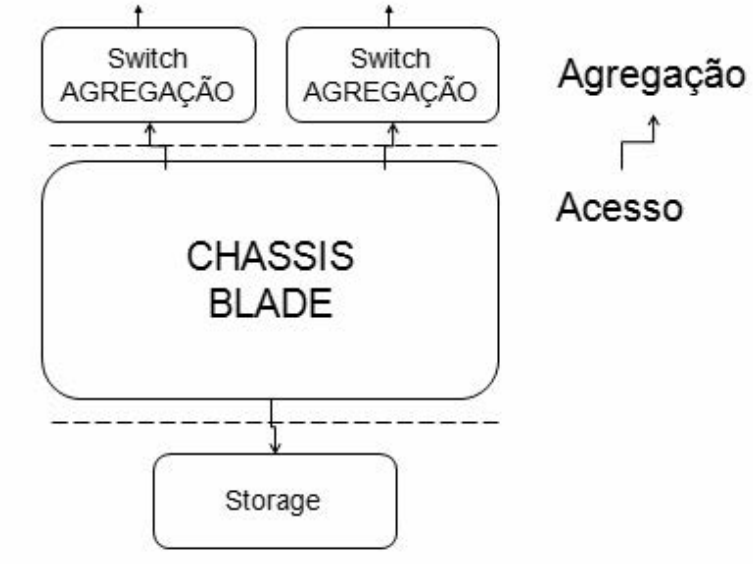
é realizado o licenciamento da aplicação e as funcionalidades obtidas com ou outro modelo. De cara, trocar um servidor de quatro processadores por dois de dois processadores é uma vantagem, pois a disponibilidade do ambiente na segunda opção é sempre maior, devido ao uso de recursos como o HA (*high availability*) do VMware vSphere.

Os servidores respondem diretamente pela VIRTUALIZAÇÃO, pois os sistemas operacionais de VIRTUALIZAÇÃO são ali instalados. O ideal é que o software de VIRTUALIZAÇÃO já seja adquirido junto com o servidor físico.

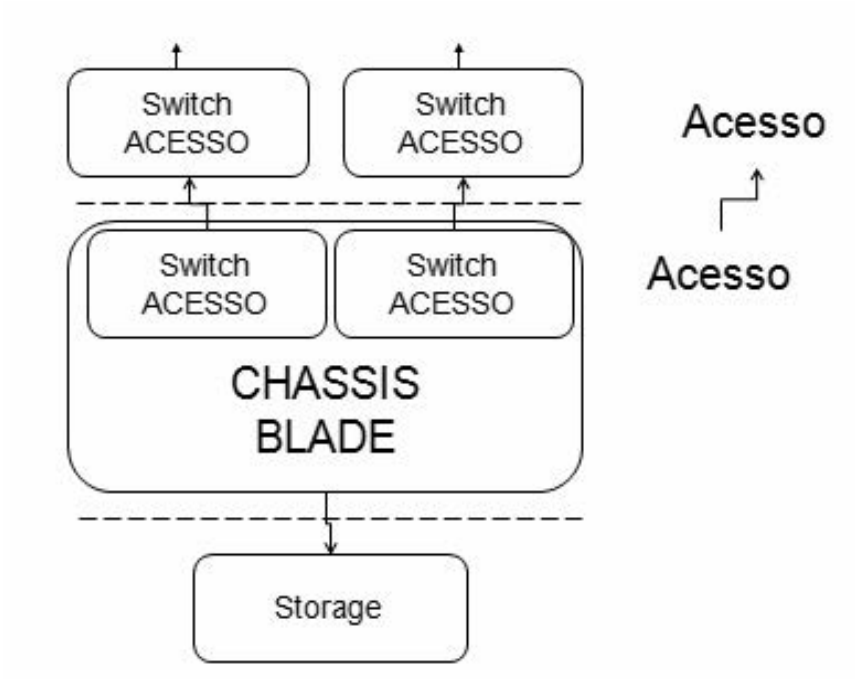
Os servidores do tipo RACK e os servidores do tipo blades são adequados para a VIRTUALIZAÇÃO. Os servidores blades especificamente mudam o perfil da densidade de energia dentro do DATACENTER e acarretam mudança na estratégia de como energizar e refrigerar o DATACENTER. Os servidores blades simplificam o gerenciamento do DATACENTER reduzindo a utilização do espaço físico e a integração de diversos componentes dentro do chassi presentes na arquitetura. Vale salientar que cada fabricante possui seu próprio chassi.

Os blades são novas estruturas de processamento, integradas em um chassi, compostos por vários servidores, as ditas lâminas, e os componentes de rede, incluindo switches de rede LAN e switches de rede SAN. O gerenciamento e os dispositivos de energia e ventilação são também integrados.

A arquitetura baseada em blades modifica a camada de acesso de rede trazendo mais um switch opcional localizado dentro do chassi. Os switches blade são funcionalmente similares aos switches de acesso e são, topologicamente falando, parte da camada de acesso, entretanto, os switches do blade são normalmente considerados como uma quarta camada no projeto da rede do DATACENTER. A **Figura 7-6** ilustra os dois casos: (a) blade com switch de agregação e (b) BLADE com switch de acesso.



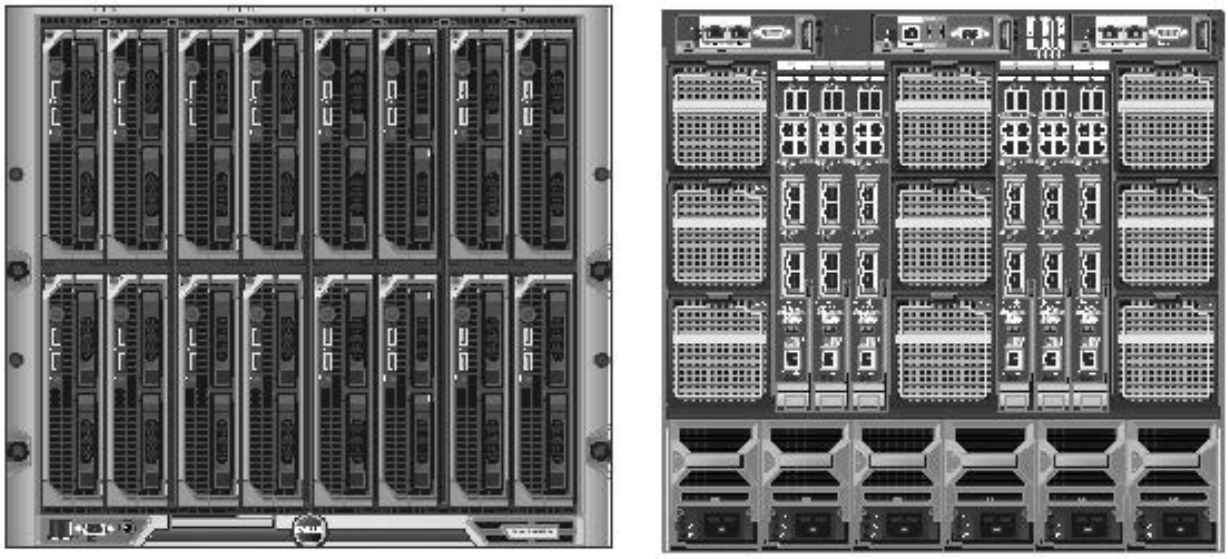
(a)



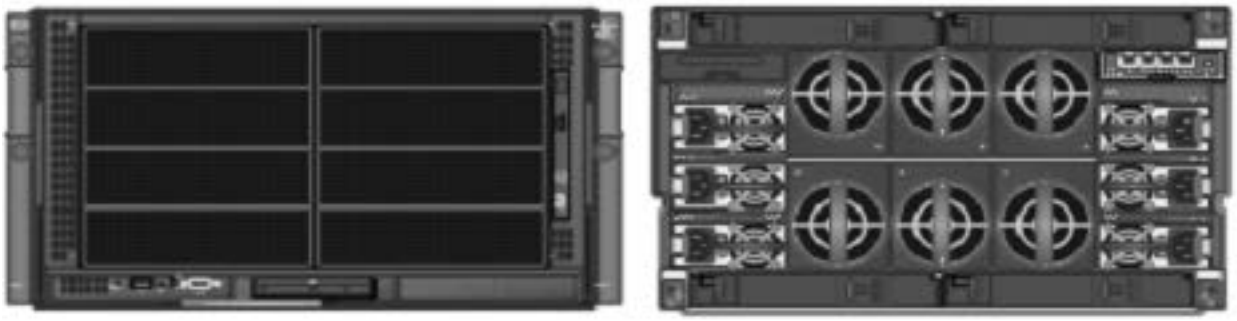
(b)

Figura 7-6 Opção de arquitetura com os BLADES

Os BLADES aperfeiçoam o uso do espaço físico, reduzem as necessidades de cabeamento, simplificam o gerenciamento e normalmente consomem menos energia quando comparados aos servidores em RACK. A altura do chassi blade típico é de 10Us. A **Figura 7-7** ilustra os BLADES da Dell (1000E) e da HP (C3000).



(a) BLADE 1000E da DELL



(b) BLADE C3000 da HP

Figura 7-7 Blades da Dell e da HP

As conexões de I/O são a parte crítica de qualquer projeto de blades, pois elas permitem a conexão do blade com o mundo externo. Em geral os blades oferecem quatro tipos de conectividade: IP/Ethernet, *Fibre Channel – FC*, *Fibre Channel over Ethernet – FCoE* e *Infiniband*.

O gerenciamento dos blades acontece em duas frentes: o gerenciamento do próprio chassi e o gerenciamento das lâminas.

7.4.2. Conceito na Prática: Dell Blade 1000E

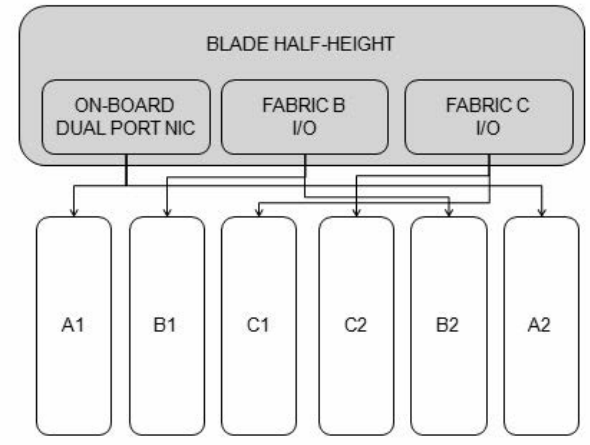
As características gerais do Blade 1000E da Dell são descritas na **Tabela 7-1**. As lâminas (blades) agora podem ser fornecidas no formato full-height.

Tabela 7-1 Blade DELL 1000E (Fonte: DELL)

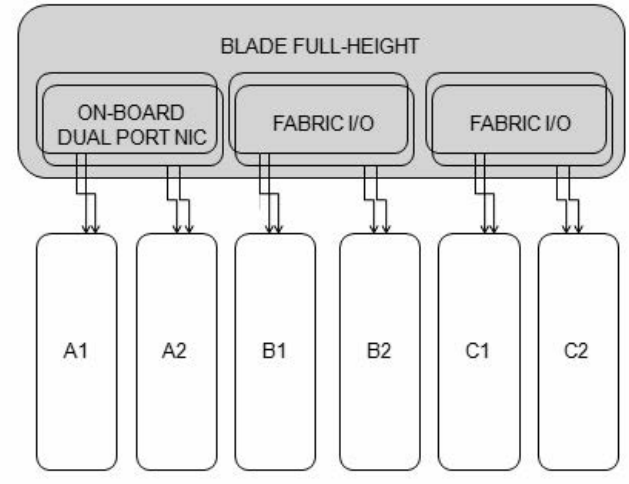
CARACTERÍSTICA	CONFIGURAÇÃO
Tamanho	**10 Us**
Blades por Chassis	**16 servidores half-height ou 8 servidores full-height**
Total de blades por RACK	**64 blades half-height ou 32 blades full-height**
Módulos de I/0	**6 (3 fabrics duais)**
Fontes de Alimentação (208 V AC)	**6 fontes em várias configurações**
Ventiladores	**9 ventiladores obrigatórios (8 + 1 redundante);**
Módulos de Gerenciamento e Interface	**2 gerenciadores de chassi CMCs (1+1 redundante); 1 iKVM; 1 painel frontal com LCD.**

As opções de I/O são diferentes para lâminas *half-height* e *full-height,* conforme pode ser observado na **Figura 7-8**. No caso (a), as lâminas são M600, M605 e M610, *half-height* com seis portas ao todo. No caso (b), as lâminas são M710, M805 e M905, *full-height* e, neste caso, existem doze portas disponíveis. Para VIRTUALIZAÇÃO a opção com mais portas (12 portas) permite otimizar o desempenho da solução como um todo, pois distribui os recursos de I/O. Pode-se distribuir melhor as redes de máquinas virtuais, movimentação de máquinas virtuais, LAN, gerenciamento, etc.

O Blade 1000E possui três *fabrics* referenciados como A, B e C. Cada fabric tem dois módulos de I/O (A1, A2, B1, B2, C1, C2) para um total de seis módulos de I/O no chassi. Cada módulo de I/O pode ser um switch ou um *pass-through* baseado em Ethernet, FC ou Infiniband. Cada *half* blade possui uma NIC dual port e dois opcionais cartões mezaninos. Um cartão mezanino é conectado ao fabric B e o outro conectado ao fabric C.



(a)



(b)

Figura 7-8 Conexões de I/O do Blade Dell 1000E (Fonte: Dell)

O tráfego gerado pela VIRTUALIZAÇÃO com um software de VIRTUALIZAÇÃO como o vSphere pode ser de quatro tipos: tráfego de máquina virtual, tráfego de gerenciamento, tráfego do VMotion (permite migração online das máquinas virtuais entre servidores físicos) e tráfego iSCSI.

Pode-se criar duas redes separadas: uma para a LAN, que inclui o gerenciamento, tráfego de máquina virtual e tráfego VMotion. Duas portas de rede em 10GB podem ser conectadas aos módulos B1 e C1, usados para a LAN. A outra rede pode ser utilizada para o tráfego iSCSI com duas portas em 10Gb conectadas aos módulos B2 e C2. A razão de fazer esta seleção é que, desta forma, obtém-se redundância através dos cartões mezaninos. Se um mezanino falhar, as redes SAN e LAN continuam a operar.

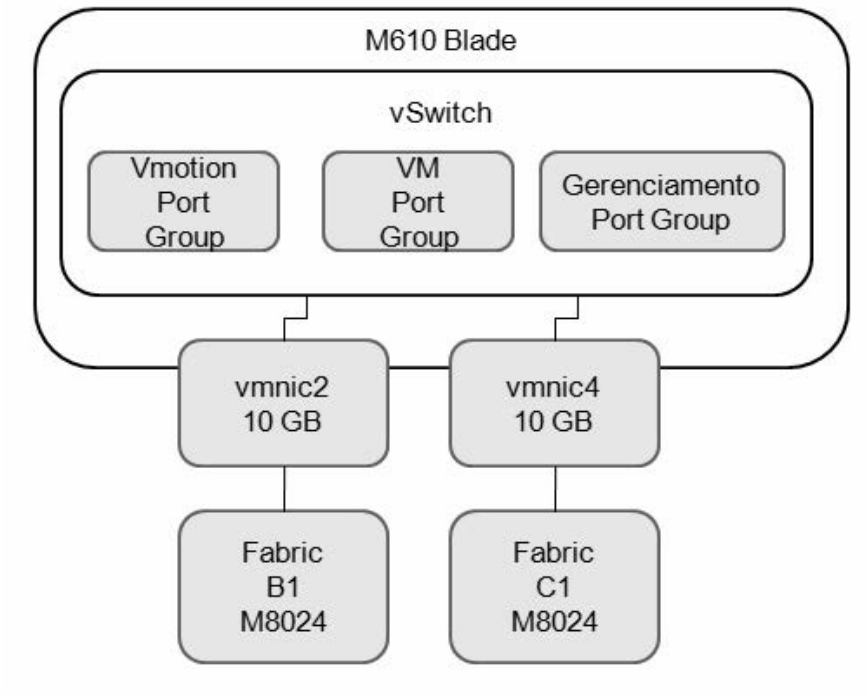
Em um projeto da arquitetura de rede pode-se adotar como princípio:

- Redundância em ambas LAN e SAN com módulos de I/O redundantes.
- Isolamento físico da rede iSCSI da LAN.
- Isolamento lógico do VMotion usando VLAN, pois o tráfego VMotion não é criptografado.
- Desempenho otimizado com balanceamento de carga.

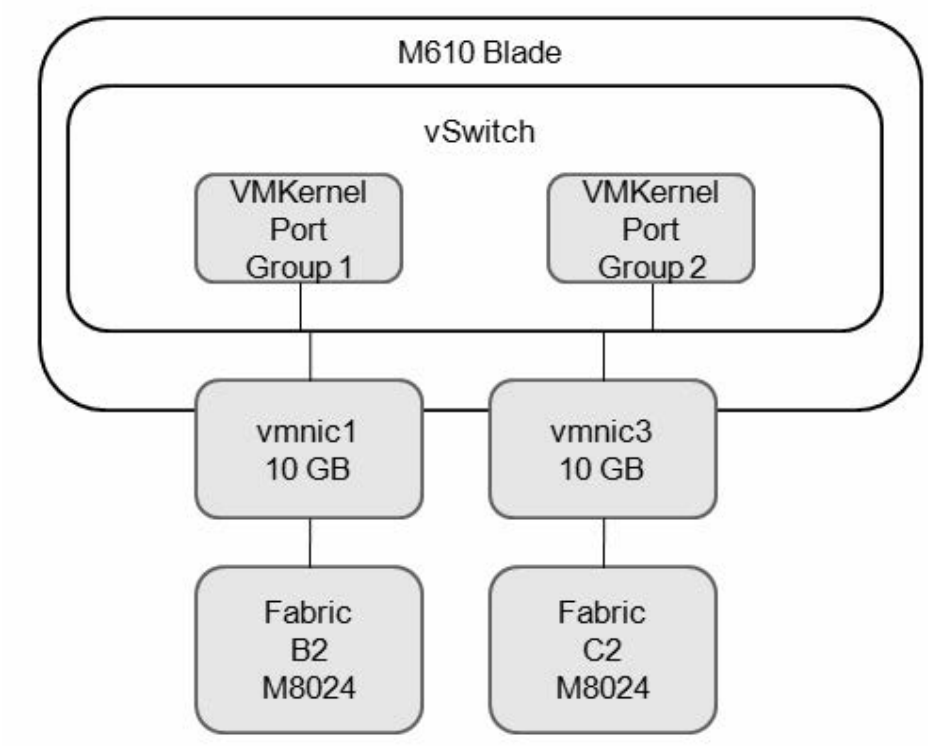
A LAN inclui os tráfegos de VM, gerenciamento e tráfego VMotion. Pode-se criar uma switch virtual (vSwitch0) e conectar duas portas de rede 10Gb físicas que funcionam como uplinks a este switch. Isto cria um time com duas portas de rede que habilita o *failover* e o *load balancing*. Estas portas de rede podem ser conectadas aos PowerConecct M8024 que ocupam os módulos B1 e C1. Três port groups são criados no vSwitch0: um para a rede VM, outro para a rede VMotion e outro para o ESX (service console) ou gerenciamento (ESXi). Quando as VMs são criadas, os adaptadores de rede virtuais são conectados ao port group da máquina virtual.

A SAN inclui o tráfego iSCSI. Um switch virtual (vSwitch1) pode ser criado e dois adaptadores de rede físicos são conectados como uplinks ao switch virtual. Estas portas dos adaptadores de rede podem ser conectadas aos PowerConnect M8024 que ocupam os módulos B2 e C2. Uma interface VMKernel é criada e associada com uma interface física de rede e utilizada como iSCSI initiator para a conexão ao storage array da SAN.

A **Figura 7-9 (a) e (b)** ilustra as duas situações.



(a)



(b)

Figura 7-9 Configuração de LAN (a) e rede iSCSI (b) com VIRTUALIZAÇÃO

O Blade M1000E da Dell possui uma patente pendente chamada de ***FlexAddress***. O ***FlexAddress*** fornece identidades de rede e de armazenamento persistentes, equipando o DATACENTER para tratar de alterações previsíveis ou, até mesmo, não planejadas – permitindo aumentar, atualizar ou remover servidores sem afetar a rede. A tecnologia ***FlexAddress*** permite que qualquer gabinete blade M-Series bloqueie o World Wide Name (WWN) do controlador de Fibre Channel e o Controle de Acesso à Mídia (MAC) do controlador iSCSI e Ethernet em um slot do blade, e não no hardware do blade como acontecia antigamente. Ao remover a identidade da rede e do armazenamento do hardware do servidor, as empresas podem atualizar e substituir componentes ou o servidor inteiro sem alterar a identidade na rede. O ***FlexAddress*** pode ser implementado em quatro etapas simples, bastando selecionar os slots e as estruturas em que se deseja que ele seja ativado e o controlador de gerenciamento de chassi cuida de todo o gerenciamento, independentemente dos módulos de I/O escolhidos.

Características Essenciais de Blades e Servidores

- **Disponibilidade:** os servidores de uma forma geral melhoraram a disponibilidade com o uso de fontes e ventiladores redundantes e discos em configuração de RAID (Redundant Array of Inexpensive Drives). Diversas pesquisas efetuadas por fabricantes demonstram que estes dois aspectos são críticos. Placas de rede duplicadas e mesmo a utilização de mais de um processador no servidor aumentam ainda mais o nível de disponibilidade. As memórias também permitem hoje a utilização de configurações redundantes. A alta disponibilidade do servidor é obtida com a configuração em cluster de pelo menos dois servidores.
- **Desempenho:** o desempenho de servidores x86 tem avançado rapidamente. Atualmente existe também a possibilidade de utilizar processadores mais econômicos do ponto de vista energético, em detrimento do desempenho. Uma comparação rápida de resultados obtidos com os benchmarks SPEC, ou mesmo o TPC, indica claramente o ganho de desempenho conseguido a cada nova atualização dos processadores.
- **Gerenciamento:** o gerenciamento do hardware do servidor é realizado através de um software de gerenciamento normalmente fornecido pelo próprio fabricante do servidor, cujo console é baseado na web. O software normalmente é instalado em uma estação de gerenciamento e utiliza o protocolo SNMP (Simple Network Management Protocol). O servidor é o cliente no protocolo SNMP. Os principais comandos são GET, SET e TRAP. Também é possível gerenciar o servidor através de uma placa de gerenciamento integrada que permite o acesso remoto ao servidor.
- **Escalabilidade:** escalabilidade é a habilidade de um sistema computacional de lidar, de forma transparente, com um número crescente de usuários ao mesmo tempo. A escalabilidade em servidores é conseguida com a adição de mais processadores. Os métodos tradicionais de escalabilidade em servidores são a esca-

labilidade vertical (scale-up) e a escalabilidade horizontal (scale-out). Em sistemas x86 é mais comum a escalabilidade do tipo scale-out utilizando um software de clustering. Em sistemas RISC é padrão a escalabilidade do tipo scale-up. O licenciamento da camada de software é sempre um aspecto importante a ser considerado na decisão por uma ou outra forma de crescimento do servidor.

- Scale-out (escalabilidade horizontal utiliza múltiplos servidores): um sistema escalável horizontal significa adicionar mais nós ao sistema para crescimento, tais como adicionar um novo servidor a um sistema de banco de dados em cluster. O scale-out é uma arquitetura de software otimizada pelo hardware. Normalmente esta modalidade apresenta menor investimento inicial e é mais flexível à introdução de novas tecnologias.
- Scale-up (escalabilidade vertical utiliza múltiplos processadores): um sistema escalável vertical significa adicionar recursos em um único nó do sistema (mais memória ou mais processamento, por exemplo). Scale-up é uma arquitetura de hardware otimizada pelo software. O principal elemento de viabilização desta modalidade de escalabilidade é o multiprocessamento simétrico (SMP) que, em certas situações e dependendo da arquitetura da aplicação, apresenta limites de linearidade. Normalmente esta modalidade apresenta maior investimento inicial e menor flexibilidade para introdução de novas tecnologias.

7.4.3. Processadores

O elemento central de um servidor é o processador. Cada processador possui um padrão determinado pelo conjunto de instruções de máquina (*Instruction Set Architecture – ISA*). Intel e AMD utilizam o padrão x86.

Os processadores Intel e AMD suportam a VIRTUALIZAÇÃO no próprio chip, diminuindo o overhead do HYPERVISOR quando do acesso ao hardware. Recentemente as novas versões de processadores AMD e Intel incorporaram as técnicas RVI e EPT, que incorporaram a VIRTUALIZAÇÃO MMU chamada Extended Page Tables (EPT), pela Intel, e Rapid Virtualization Index (RVI), pela AMD. Esta modalidade de VIRTUALIZAÇÃO no nível do hardware elimina a necessidade de utilizar shadow pages tables.

A computação baseada no padrão x86 vem ganhando escala e a demanda por maior desempenho é crescente. A arquitetura x86 (x86 se refere a instruções e registradores utilizados) continua a evoluir, mas mantém instruções derivadas da arquitetura 8086 de 16 bits. Este talvez seja o grande segredo da aceitação desta arquitetura.

Os fabricantes de processadores x86 mantêm a compatibilidade de instruções para trás e para frente com o padrão. Existem duas principais arquiteturas x86, baseadas em 32 e 64 bits. A Intel convencionou chamar de (IA) 32 e (IA) 64 para as versões de 32 bits e 64 bits da arquitetura x86. Por sua vez, a microarquitetura refere-se ao projeto físico de cada processador. Processadores com diferentes microarquiteturas ainda assim podem utilizar um mesmo conjunto de instruções, ou seja, a mesma arquitetura, que é o que acontece com os processadores Intel e AMD. A microarquitetura Intel evoluiu passando da opção Netburst para Core e, agora, para o Nehalen. Os processadores AMD

OPTERON são também largamente utilizados na indústria de servidores e apresentam excelente desempenho.

Uma ideia importante em servidores é a de utilizar o conceito de benchmark. O benchmark permite comparar duas CPUs de fabricantes distintos para um mesmo setup e condições de carga ou mesmo comparar os resultados obtidos por CPUs diferentes de um mesmo fabricante. Existem diversos benchmarks para diversas funções executadas pelos servidores. O benchmarking mais simples é o orientado para o throughput da CPU, como o SPEC CPU 2006, cujos resultados para as CPUs de mercado podem ser obtidos no site do SPEC (*Standard Performance Evaluation Corporation*, www.spec.org).

Existem também os benchmarks para processamento de transações, que medem a habilidade de um sistema para tratar transações. Estes benchmarks consistem em verificar o acesso a sistemas de banco de dados e atualizações. Na década de 80, um grupo de engenheiros criou o TPC (*Transaction Processing Council, www.tpc.org*), uma organização independente de fornecedores com o intuito de criar benchmarks equilibrados e realistas para o processamento das transações. O TPC-C, orientado para o benchmark de sistemas gerenciadores de banco de dados sob consultas complexas, foi criado em 1992 e hoje está na versão 5.10.1.

A utilização de servidores baseados na plataforma x86 tomou impulso com o avanço da arquitetura cliente/servidor no início da década de 90. Os servidores foram aprimorados em termos de confiabilidade, e o poder de processamento aumentou com o rápido avanço dos processadores, possibilitando o seu uso em quase todo tipo de aplicação. Ainda existem limitações para determinadas *cargas de trabalho* mas, de uma forma geral, os servidores x86 atendem às demandas do mundo corporativo.

7.4.4. Conceito na Prática: Intel Processador XEON 5600

O novo processador Intel 5600 foi construído com foco em aumento de desempenho para sistemas virtualizados, sem descuidar da eficiência energética. Em sua configuração máxima, são seis cores em um projeto de 32nm. O desempenho é cerca de 60% superior, quando comparado à família anterior 5500, e o consumo é cerca de 30% inferior, quando comparado com a mesma família. As opções são por seis, quatro ou dois cores. A **Figura 7-10** ilustra a arquitetura do Intel XEON 5600.

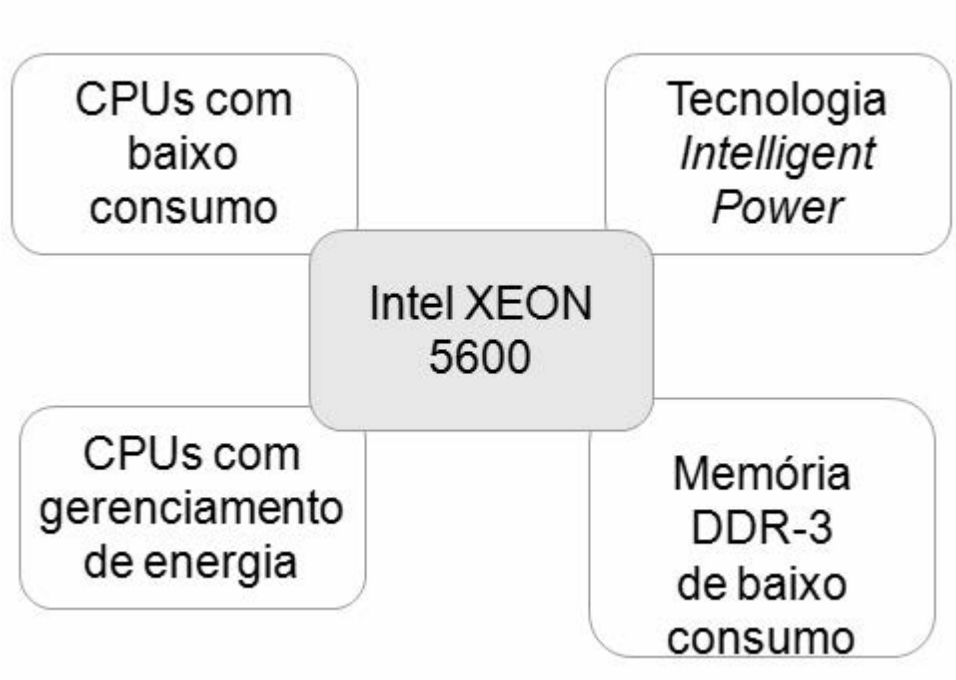

CPUs com baixo consumo
Tecnologia *Intelligent Power*
Intel XEON 5600
CPUs com gerenciamento de energia
Memória DDR-3 de baixo consumo

Figura 7-10 Intel XEON 5600

O Intel 5600 possui três grandes famílias com foco em desempenho, preço versus desempenho ou custo. Para máximo desempenho, as frequências de trabalho são maiores e o consumo de energia também. Para baixo custo, as frequências de trabalho são menores e podem ser adquiridas com quatro ou dois cores. A **Figura 7-11** ilustra as três opções com foco em:

a) Desempenho.

b) Preço versus Desempenho.

c) Custo.

Estas opções existem, logicamente, porque as demandas são diferentes. Existem aplicações cujas demandas são de alto desempenho e outras demandam baixo custo em detrimento do desempenho. Também existem alternativas de melhor custo desempenho. Também, dentro das três alternativas, pode-se otimizar o consumo ou a máxima frequência de operação. Em um DATACENTER com 1000 processadores XEON 5600, o consumo do DATACENTER pode variar de 40KW até 130KW. Ou seja, ganha-se em desempenho, mas paga-se mais pela energia.

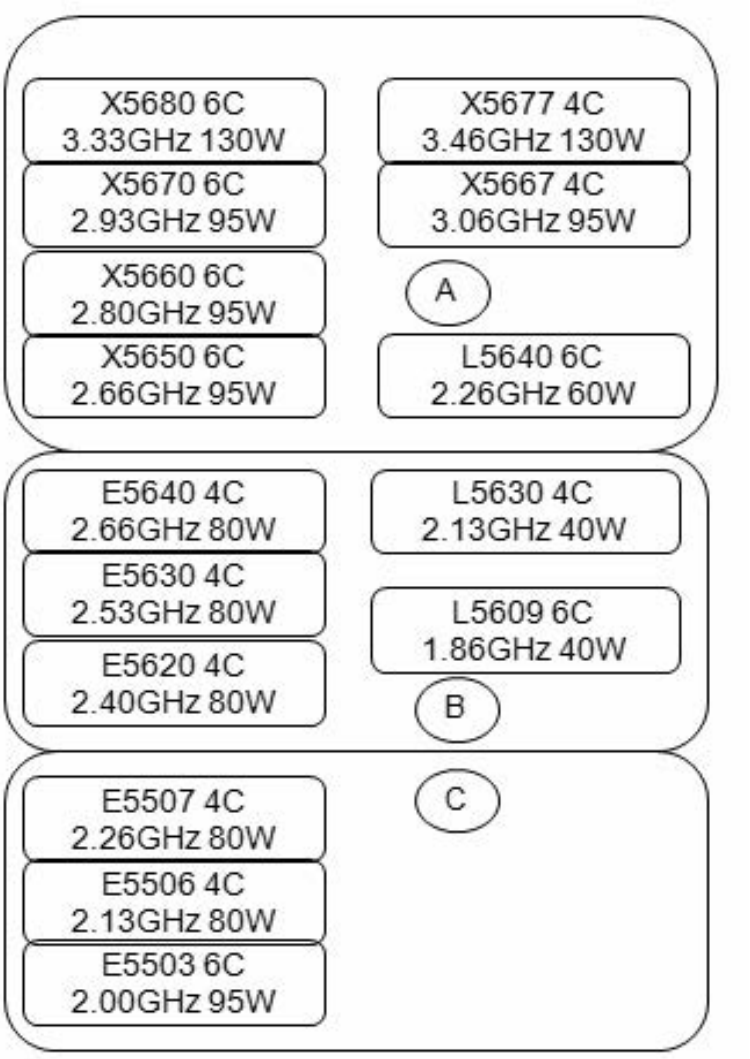


Figura 7-11 Famílias do XEON 5600

O Processador Intel XEON 5600 assiste de forma robusta a VIRTUALIZAÇÃO e expande a segurança com novas funcionalidades como AES-NI e Intel TXT. Intel VT & TXT isolam as VMs e aumentam a segurança da plataforma; Intel TXT estabelece o *status trusted* que habilita a migração baseada em políticas de segurança. INTEL ES-NI permite encriptação acelerada que melhora a proteção dos dados.

7.4.5. Licenciamento

Servidores com processadores com vários núcleos e VIRTUALIZAÇÃO aumentam a probabilidade do licenciamento do software utilizado estar inadequado.

A maior parte dos softwares para servidor ainda está licenciada por soquete (CPU). O raciocínio é simples: os chips são fáceis de serem mensurados e essa maneira de obter a licença mobiliza o pessoal de TI para sempre utilizar chips mais poderosos.

Obter o máximo de um software sempre exigiu hardware de alto desempenho, a diferença é que atualmente o desempenho aumenta com o aumento da quantidade de núcleos, em vez de aumentar com a quantidade de megahertz (MHz). Esta estratégia da indústria está ligada ao fato de que o crescimento através do aumento de núcleos é mais interessante, pois reduz o consumo de energia. Além disso, a linearidade que eventualmente poderia ser perdida com a utilização de muitos núcleos tem sido otimizada pelos fabricantes (Intel e AMD).

Quando se trata do licenciamento de servidores multicore é preciso considerar dois aspectos novos: a quantidade de cores e a VIRTUALIZAÇÃO. A política de licenciamento para fabricantes diferentes não é a mesma.

- **VMware** & **Citrix**: a VMware especificamente adotou o licenciamento por soquete, assim como fez a Citrix com o XenServer.
- **Microsoft**: a Microsoft trata cada máquina virtual como um servidor físico com o mesmo número de soquetes do servidor real. As versões do Windows Server 2008 incluem licenças para instâncias virtuais extras do software na mesma CPU – uma na versão Standard Edition; quatro na versão Enterprise Edition e número ilimitado na versão Datacenter Edition. Todas as licenças para servidor da Microsoft também incluem direitos de downgrade, o que significa que uma instância virtual poderá ser substituída pelo Windows 2000 ou Windows NT. A Microsoft não faz diferença na forma de licenciamento entre tecnologias ou fabricantes de VIRTUALIZAÇÃO.

Os direitos de virtualização por licenciamento de servidor, quando do uso do Hyper-V, precisam ser considerados e variam de acordo com a versão do Windows Server utilizada. A **Tabela 7-2** ilustra os requisitos de licença por servidor para o Hyper-V.

Tabela 7-2 Requisitos de Licença para o Hyper-V

Hyper-V Server 2008	Windows Server 2008 Standard	Windows Server 2008 Enterprise	Windows Server 2008 DATACENTER
VM requer uma licença de servidor	1 máquina física + 1 VM	1 máquina física + 4 VMs	1 máquina física + número ilimitado de VMs

A Microsoft recomenda que o Hyper-V seja o único perfil habilitado em servidores utilizados na virtualização.

- **IBM**: a IBM conta núcleo de CPU levando em consideração o desempenho de cada linha de processador. Atualmente, o licenciamento por núcleo da IBM se aplica a quase metade de seus produtos de software, incluindo DB2, WebSphere, Tivoli e Domino. O esquema da IBM é o mais complexo, sem dúvida. Ela cobra dos clientes por PVU (*processor value unit*, ou unidade de valor por processador), que é igual a 1% do preço de um processador padrão com núcleo duplo, Opteron ou Xeon. A IBM dividiu seus preços por soquete por 100 para definir o preço por PVU, o que significa que a maioria dos clientes x86 inicialmente não verá nenhu-

ma mudança. Contudo, ao fazer a atualização para os chips x86 com processadores quadcore, o usuário pagará o dobro por processador, porque cada um deles equivale a 200 PVUs.

A IBM afirma que este modelo é justo porque um chip com quatro núcleos pode fazer o dobro do que faz um chip com dois núcleos. O problema para o pessoal de infraestrutura de TI é que a quantidade de núcleos das CPUs está aumentando. Isso significa que ocorrerá o mesmo com as taxas de licenciamento da IBM, se o número de PVUs por núcleo e o preço por PVU permanecerem constantes. Como esta situação é insustentável em longo prazo, a IBM declara que ajustará o número de PVUs por núcleo para dar conta do verdadeiro desempenho e espera diminuir os custos de acrescentar mais núcleos a um chip. Em 2008, todos os núcleos x86 ainda respondem por 50 PVUs.

O sistema de PVU, da IBM, tem um aspecto interessante por levar em conta, explicitamente, a VIRTUALIZAÇÃO, por meio do que a IBM chama de "licenciamento de subcapacidade". Se um servidor for dividido em diversas máquinas virtuais, os aplicativos dentro de cada uma delas precisarão ser licenciados somente para o número máximo de cores disponíveis para cada VM, e não para todos os núcleos no servidor.

- **ORACLE**: a ORACLE conta núcleo de CPU levando em consideração o desempenho de cada linha de processador de uma forma diferente da IBM. O licenciamento da Oracle por processador (por usuário não se aplica aqui) é mais simples do que o da IBM, sendo efetuado com base na contagem de núcleos como frações de um processador, mas é menos favorável à VIRTUALIZAÇÃO. Um banco de dados Oracle sendo executado em VMware deve ser licenciado para cada núcleo no hardware básico – independentemente de em quantos núcleos a VM realmente é executada.

Atualmente, o único meio de economizar em licenciamento da Oracle por meio da VIRTUALIZAÇÃO é limitar os núcleos do processador disponíveis para um banco de dados. Por isto que a Intel e a AMD continuam fabricando processadores XEON dual core, que são comercializados pelo mesmo valor dos Xeon edition quadcore.

A Oracle para a versão do banco de dados Enterprise Edition considera o total de núcleos do servidor com um fator de multiplicação de 0.5 para processadores Intel e AMD x86. um exemplo prático: um servidor com quatro processadores six-core possui um total de 24 cores. Número de licenças requeridas por processador=Total#CORES rodando Oracle x 0,5. Para efeito de licenciamento a conta seria a seguinte: 24 x 0.5 = 12 licenças. No caso da versão Standard Edition, a ORACLE considera para efeito de licenciamento a quantidade de sockets, mas limita o uso da versão em servidores de até 4 soquetes.

7.4.6. Conceito na Prática: Calculadoras do Windows Server 2008

As calculadoras disponibilizadas pela Microsoft no seu site facilitam o cálculo das licenças para ambientes virtualizados. As calculadoras de VIRTUALIZAÇÃO do Windows Server 2008 R2 propiciam dois caminhos para estimar o número e o custo das licenças de Windows Server Standard, Enterprise e DATACENTER, necessárias para cenários de VIRTUALIZAÇÃO, e ajuda assim a determinar o melhor custo efetivo para as edições do Windows Server 2008.

A calculadora de VIRTUALIZAÇÃO 1, mostrada na **Figura 7-12,** é utilizada para estimar as licenças e o custo das edições do Windows Server e produtos baseados no Windows Sever em um único servidor físico. A ferramenta de VIRTUALIZAÇÃO não precisa ser obrigatoriamente da Microsoft para utilizar a calculadora.

A calculadora permite construir máquinas virtuais rodando múltiplos produtos Microsoft para estimar as licenças e custos do Windows Server por edição (Standard, Enterprise e Datacenter) e muitos produtos Microsoft baseados no servidor. Inicia com a definição da configuração para a tecnologia de VIRTUALIZAÇÃO a ser utilizada, soquetes e requisitos de cluster e constrói as máquinas virtuais.

Figura 7-12 Calculadora de VIRTUALIZAÇÃO 1

A calculadora de VIRTUALIZAÇÃO 2, mostrada na **Figura 7-13,** é utilizada para estimar as licenças e o custo por edição do servidor Windows para um ou múltiplos servidores físicos.

A calculadora 2 propicia dois caminhos para estimar o número e o custo das licenças necessárias de Windows Server Standard Edition, Enterprise Edition e Datacenter Edition para os cenários de VIRTUALIZAÇÃO e ajuda a determinar assim o melhor custo efetivo da edição do Windows Server. Pode-se utilizar dois modelos para o cálculo: agrupar servidores pelo número de processadores e pela média de VMs, ou listar cada servidor individualmente.

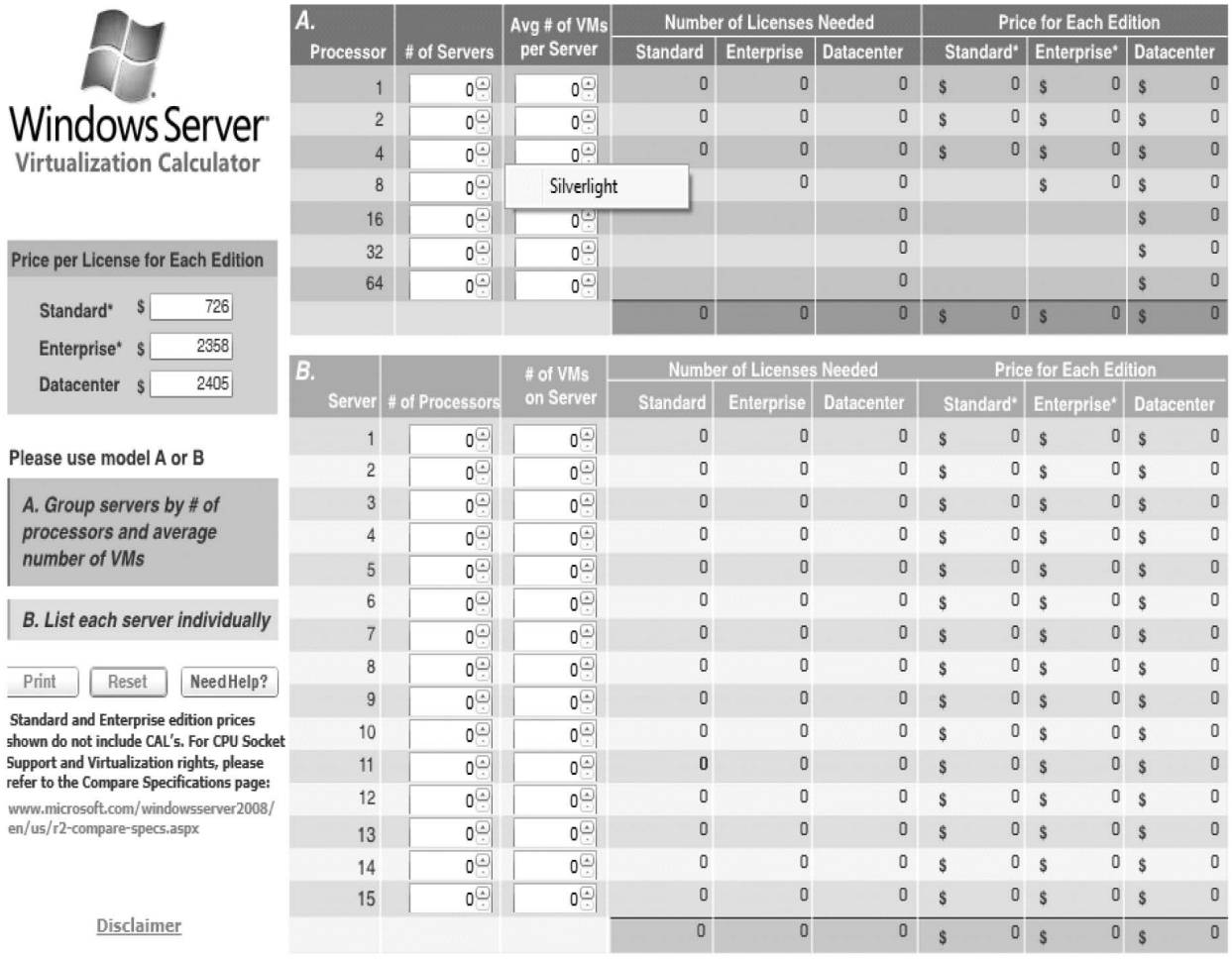


Figura 7-13 Calculadora de VIRTUALIZAÇÃO 2

7.5. Rede e VIRTUALIZAÇÃO

O ideal em um projeto de DATACENTER virtualizado é que as políticas de rede sejam mantidas, mesmo que as máquinas virtuais se movimentem, isto é, as máquinas virtuais e suas políticas devem ser independentes da localização física.

A **Figura 7-14** ilustra as opções de conectividade na camada de acesso para blades, virtualizados ou não. Observe que, no caso dos blades com VIRTUALIZAÇÃO, os switches aparecem em vários níveis, tornando a camada de acesso complexa e de difícil gerenciamento.

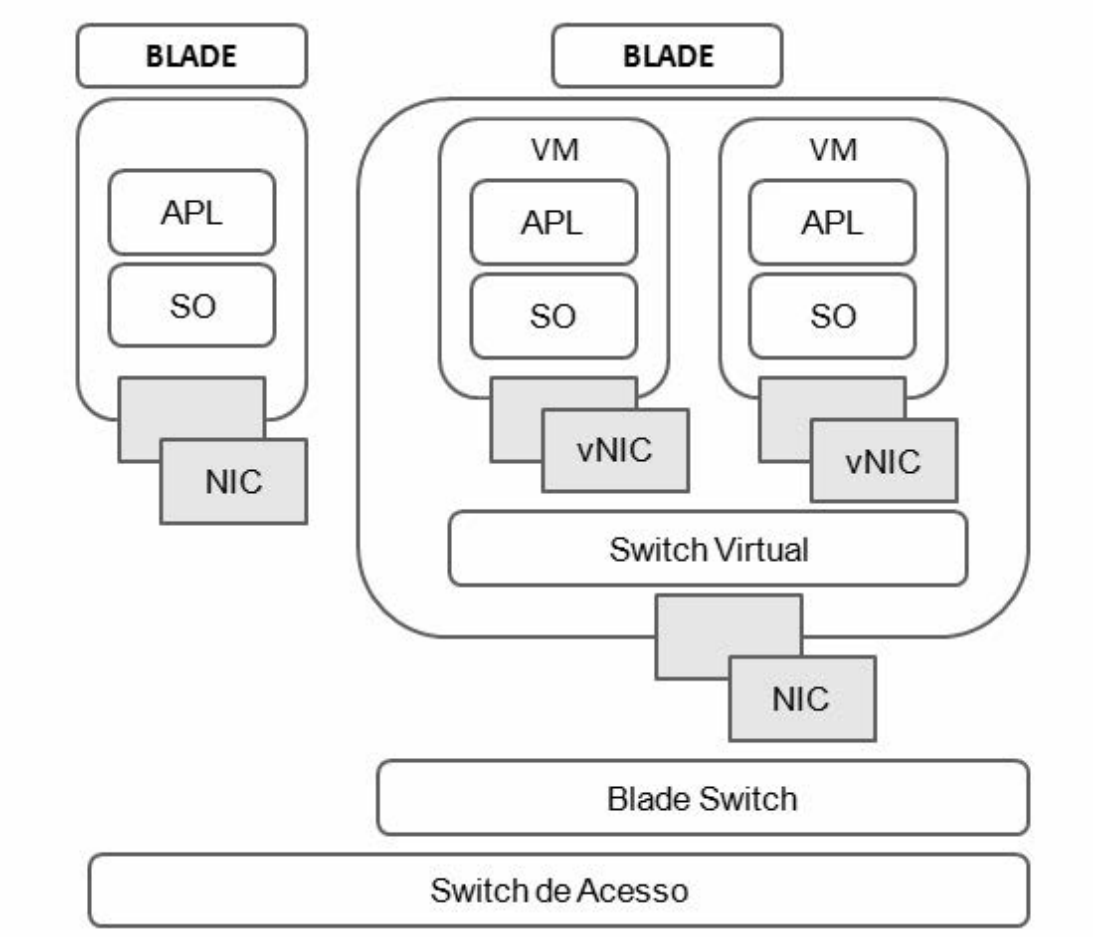


Figura 7-14 Opções para a camada de acesso

Os fabricantes de componentes de rede e os fabricantes do software de VIRTUALIZAÇÃO passaram a trabalhar na redução desta complexidade. A VMware resolveu estas dificuldades implementando um switch por software (vSwitch) como parte do HYPERVISOR. Cada placa de rede virtual (vNIC) conecta-se logicamente a máquina virtual ao vSwitch e possibilita que a máquina virtual envie e receba o tráfego nesta interface. Se duas vNICs conectadas ao mesmo vSwitch precisam se comunicar, o vSwitch é o próprio switch de acesso e não precisa enviar o tráfego para a rede física. Esta opção é simples, mas traz uma série de inconvenientes, dificultando o gerenciamento da rede. Neste *approach* cada vSwitch passa a ser mais um ponto de configuração.

A VMware e a Cisco criaram o conceito de switch virtual distribuído (DVS) que possibilita que os diversos vSwitches possam ser gerenciados de um único ponto. A VMware implementou o DVS como vNetwork Distributed Switch e o controle é realizado pelo VMware vCenter. A Cisco chama esta funcionalidade de Cisco Virtual Network Link (VN-Link) . O termo indica a possibilidade de criação de links lógicos entre uma vNIC em uma máquina virtual e um switch Cisco com o VN-Link habilitado. A base do conceito é a criação da interface Ethernet virtual (vEth).

As interfaces vEth são o equivalente das portas de acesso da rede física. Um switch habilitado para VN-Link pode implementar várias interfaces vEth por porta física e criar um mapeamento entre cada vEth e o correspondente vNIC na máquina virtual. A **Figura 7-15** ilustra as relações entre redes físicas e virtuais em um switch com o VN-Link habilitado.

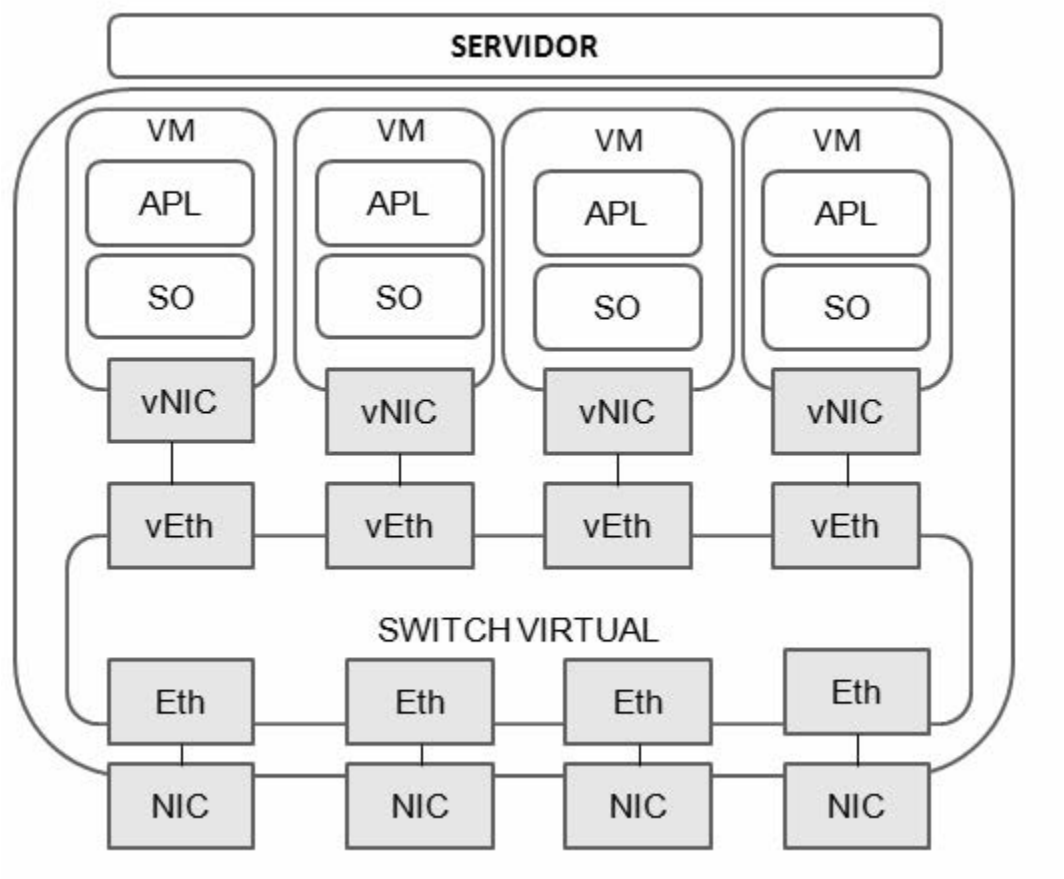


Figura 7-15 Redes Físicas e Virtuais com o VN-Link

A tecnologia VN-Link pode ser implementada por dois caminhos:

- Cisco DVS rodando dentro da camada HYPERVISOR (Cisco Nexus 1000V).
- Como uma nova classe de dispositivos que suportam a VIRTUALIZAÇÃO da interface de rede (NIV-network interface virtualization), eliminando a necessidade de switches baseados em software dentro dos HYPERVISORS.

7.5.1. Conceito na Prática: Cisco Nexus 1000V

Cisco Nexus 1000V é um switch virtual implementado por software em ambiente VMware vSphere. O Cisco Nexus 1000V opera dentro do HYPERVISOR da VMware e suporta a tecnologia VN-Link.

O switch Nexus 1000V foi desenvolvido em parceria com a VMware e permite que a máquina virtual tenha consistência de rede mantendo a configuração, as políticas de

segurança, as ferramentas de diagnóstico e os modelos operacionais dos servidores físicos.

O switch Nexus 1000V tem dois componentes: Virtual Ethernet Module (VEM), que roda dentro do HYPERVISOR, e um módulo de supervisão externo VSM. O VEM substitui o vSwitch do VMware. O VSM controla múltiplas VEMs como um único switch lógico. O VN-Link permite utilizar um modelo único de gerenciamento dos switches físicos e virtuais.

A **Figura 7-16** ilustra a arquitetura de um DATACENTER baseado no VMware que utiliza o Nexus 1000V.

As opções de rede virtual são providas pelo software de virtualização e gerenciadas pelo seu componente de gerenciamento. Com a rede virtual é possível criar redes virtuais dentro de um servidor ou através de múltiplos servidores.

Os dois componentes de uma rede virtual são:

- **Switches virtuais**: as funcionalidades de um switch virtual são as mesmas de um switch camada 2, suportando inclusive VLANs com controle pelas portas. O protocolo spanning tree não é necessario, pois é uma arquitetrura de uma só camada.
- **NICs virtuais**: permitem conectar VMs entre si. Os NICs físicos funcionam como uplinks para as portas do switch virtual. O NIC Teaming possibilita conectar um switch virtual a múltiplos adaptadores Ethernet. As VMs podem ser configuradas com um ou mais NICs virtuais, cada um com seu endereço IP e endereço MAC.

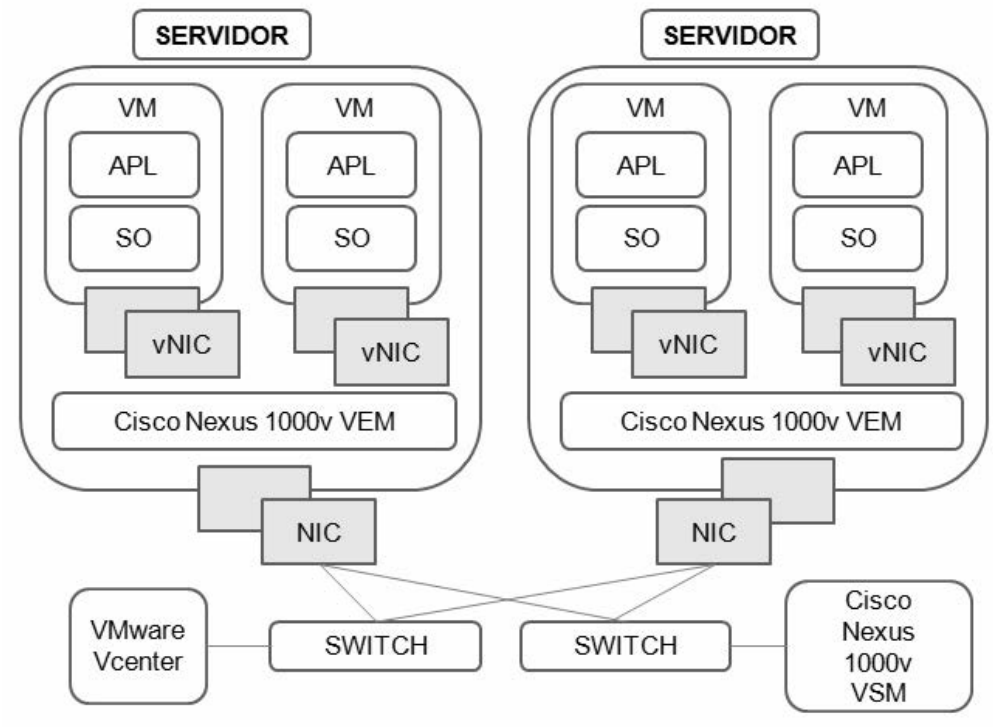


Figura 7-16 Switch Cisco Nexus 1000V

Considerações importantes sobre a adoção de VIRTUALIZAÇÃO:

- Portas NIC são determinantes de desempenho.
- O HYPERVISOR deve ter um *fabric* específico.
- Recomendação geral: ter NIC exclusiva para o gerenciamento, considerar 2 NICs se a alta disponibilidade for requerida, NICs dedicadas para funcionalidades como migração viva, 2 ou mais NICs redundantes para a rede de armazenamento, se for baseada em iSCSI, e duas ou mais NICs redundantes para as conexões das VMs.

7.6. Storage e VIRTUALIZAÇÃO

7.6.1. Introdução

O storage é hoje um componente crítico da infraestrutura de TI e é responsável direto pelo nível dos serviços de armazenamento entregues pelo DATACENTER. A necessidade de suportar o crescimento da massa de dados digital e ao mesmo tempo a necessidade de aumentar a confiabilidade dos dados, devido a aspectos regulatórios e a operações cada vez mais baseadas em 24 por 7, tornam-no o ponto focal de muitos projetos de VIRTUALIZAÇÃO da infraestrutura de TI. O storage responde pelo requisito de entrada/saída (I/O) do sistema computacional.

7.6.2. Discos e Interfaces

A entrada e saída (I/O) sempre foi negligenciada na arquitetura de um sistema de computador e hoje representa um gargalo. A evolução contínua dos processadores tem provocado o aumento do gap de desempenho existente entre processadores e sistemas de armazenamento. Por exemplo, os discos mecânicos, ainda padrão na maioria das instalações de storage, impõem limites ao desempenho do I/O devido às características mecânicas e rotacionais deste tipo de dispositivo quando em operação de leitura e escrita. Na prática o processador fica esperando a operação de I/O, prejudicando o desempenho das aplicações e do sistema como um todo.

Recentemente apareceu uma opção para os discos mecânicos (*hard disk drives – hdd*), que são os discos SOLID-STATE, ainda com custo proibitivo, mas que deverão ser rapidamente utilizados devido ao fato de melhorar significativamente o desempenho do I/O.

Os discos SSD são baseados em memória não volátil NAND e podem melhorar o TCO para aplicações intensivas em I/O, melhorar o desempenho em menos espaço, reduzir o consumo de energia e aumentar a confiabilidade.

Os discos SSD emulam os discos HDD que são manufaturados utilizando os mesmos form factors e usam interfaces como SATA (*Serial Advanced Technology Attachment*) e FC (*fibre channel*). Os discos SSD são mais baratos do que os discos HDD por IOPS.

A Intel conduziu uma prova de conceito (*proof of concept – POC*) com discos SSD e discos HDD em arrays de disco e em servidores. Os testes comprovaram que os discos SSD apresentaram melhor performance em até 8x para um TCO equivalente. Os discos

utilizados foram os X-25 Extreme SATA (64GB) contra os discos HDD de 73GB e 15000 RPM. [10]

A Intel também tem verificado a possibilidade de utilizar discos SSD como disco de swap para grandes *workloads*. Esta opção permite sustituir memória RAM de alto custo por discos SSD de menor custo. Os testes indicaram uma vantagem de 74% em termos de custo normalizado por desempenho e que a opção por fazer o SWAP em disco SSD ficou a 88% da opção de pôr a workload inteira na memória RAM.[11]

7.6.3. Sistema de Armazenamento

O storage é só um dos componentes de um sistema de armazenamento. Na verdade existem três principais componentes em um sistema de armazenamento:

- **Servidores**: os servidores são onde as aplicações rodam. Os usuários armazenam e recuperam dados através das aplicações que rodam nos servidores. Os servidores consistem de componentes físicos (hardware) e componentes lógicos (software). Os componentes lógicos do servidor são:
  - **O sistema operacional**: controla todos os aspectos do ambiente computacional.
  - ***Os device drivers:*** softwares especiais que permitem que o SO interaja com dispositivos específicos.
  - **O volume manager**: é o software que roda no servidor e faz a interface entre o sistema de arquivo e o disco físico.
  - **O sistema de arquivos**: relacionado à estrutura hierárquica de arquivos.
  - **A aplicação**: interface entre o usuário e o servidor. O acesso aos dados pode ser feito por blocos ou por arquivos. O acesso por blocos é o mecanismo básico de acesso aos discos; por sua vez, o acesso por arquivos é uma abstração do acesso por blocos.
- **Storage**: o storage é o componente principal do sistema de armazenamento. Pode utilizar um meio magnético ou de estado sólido. Discos e fitas utilizam o meio magnético. Discos ópticos utilizam um meio de estado sólido.
- **Conectividade**: refere-se à interconexão entre o servidor e o dispositivo de armazenamento, o storage. A conectividade possui componentes físicos (hardware) e componentes lógicos (protocolos).

7.6.4. Protocolo SCSI

O protocolo SCSI foi desenvolvido para propiciar um mecanismo de transporte eficiente de dados entre os servidores e os periféricos, como discos e outros recursos. A arquitetura do padrão SCSI-3 (Modelo de Arquitetura SCSI – SAM3) define um modelo cliente/servidor onde existe um initiator (servidor), um target (disco) e um subsistema

10 Solid-State Drives in the Enterprise: A Proof of Concept, Intel, março de 2009.

11 Advantages of Solid-State Drives for Design Computing, Intel, setembro de 2009.

de entrega que pode ser um cabo paralelo, Fibre Channel ou iSCSI. A arquitetura do SCSI-3 envolve os comandos SCSI-3, os protocolos de transporte e a interconexão física que possibilita transferir os dados entre o initiator e o target.

7.6.5. Características do Storage

Diferente de um servidor de arquivos, um storage pode ser visualizado como um servidor de discos. Um servidor quando conectado ao storage só enxerga os discos no storage e utiliza o sistema de arquivos fornecido pelo próprio sistema operacional.

A **Figura 7-17** ilustra a arquitetura básica de um storage com seus principais elementos.

- Portas.
- Cache.
- Controladora de Discos.
- Discos.

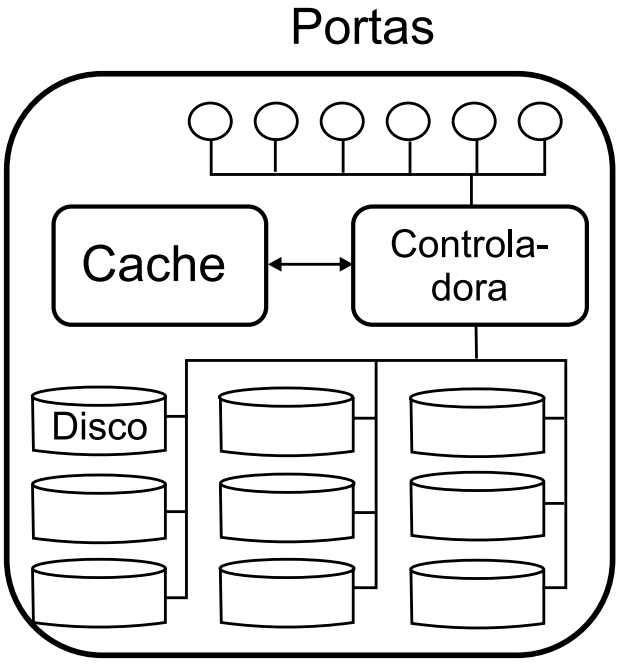


Figura 7-17 Arquitetura de um Storage

Os servidores são conectados ao storage localmente ou usando tecnologias adequadas de armazenamento em rede, como fibre channel e iSCSI, e podem utilizar a capacidade de discos propiciada pelo storage. As portas de conexão são estendidas para os discos internamente por meio de canais de I/O. Grandes subsistemas de discos podem ter várias portas de conexão, controladoras redundantes e vários canais de I/O internos e podem armazenar vários terabytes de dados. O cache da controladora desempenha papel importante no desempenho do storage.

A conexão realizada via rede aumenta a quantidade de servidores que podem ser conectados ao storage e otimiza o uso dos discos. O espaço em disco livre pode ser vinculado a qualquer servidor conectado ao storage, o que resolve o principal problema do uso local de discos onde discos subutilizados não podem ser utilizados de maneira convencional por outros servidores.

7.6.6. Tipos de Storage

Os tipos de storage podem ser divididos em três categorias: JBOD, storage do tipo RAID e storage inteligente.

- **JBOD**: se o subsistema de disco não possui controladora interna, é considerado um JBOD (*just a bunch of disks*). No caso de JBOD, as controladoras não possuem a tecnologia RAID e fazem parte do servidor que estará conectado ao storage.
- **RAID**: o storage redundante de discos de baixo custo (*redundant array of inexpensive disks* – RAID) ou para alguns autores storage redundante de discos independentes (RAID – *redundant array of independent disks*) foi uma inovação surgida em 1987, na Universidade de Berkeley, que permitiu melhorar o desempenho do sistema de I/O e, ao mesmo tempo, trouxe o aumento da disponibilidade em subsistemas de disco, um aspecto crítico em sistemas de DATACENTER. Com a utilização do RAID, também é possível aumentar o throughput do sistema de I/O com o uso de muitas unidades de disco, em vez de utilizar uma grande unidade. As controladoras RAID também utilizam o artifício da memória cache para otimizar o desempenho do subsistema de discos. As técnicas utilizadas para realizar o RAID são o espelhamento e a paridade.

Os principais níveis de RAID utilizados são 0,1,0+1,1+0,3,5,6. Quando não se consegue levantar as características de I/O de uma aplicação (% random & sequential read; % random & sequential write; tamanho do bloco, etc.), o que é muito comum, a proteção mais adequada será sempre RAID-10 ou RAID-1 para melhor desempenho, embora o custo seja sempre mais elevado.

- **Storage Inteligente**: um sistema de armazenamento inteligente possui quatro componentes chaves: front end, memória cache, back end e discos físicos.
  - **Front-end**: Propicia a interface entre o sistema de armazenamento e o servidor. Possui normalmente portas de front-end e controladoras front-end. As portas executam os protocolos apropriados.
  - **Memória cache**: nos subsistemas de disco os caches são utilizados para acelerar o acesso de leitura e escrita nos discos físicos. O cache existe no disco e nas controladoras (escrita e leitura). Todo disco possui um pequeno cache que serve para operações de escrita e leitura. A taxa de transferência do canal de I/O para a controladora é significativamente mais alta do que a velocidade que a controladora pode escrever ou ler do disco, por isso aqui também vale a ideia do cache. No subsistema de disco o cache de escrita da controladora normal-

mente é espelhado e tem uma bateria de backup que salva os dados no cache de escrita em caso de crash da controladora. A aceleração na leitura já é mais complicada quando o acesso é aleatório, pois a controladora não sabe exatamente onde estará a nova demanda proveniente da aplicação.

- **Back-end**: propicia uma interface entre a memória cache e os discos físicos. Consiste de dois componentes: portas back-end e controladoras back-end.
- **Discos Físicos**: os discos físicos são conectados ao *back-end* com interface SCSI ou Fibre Channel. O storage inteligente permite o uso de discos misturados como SCSI, FC ou SATA.

Em geral os sistemas inteligentes possuem, além do RAID e do cache, quatro principais funcionalidades: permitem realizar cópias instantâneas, permitem fazer o espelhamento remoto, oferecem o *lun masking* que controla o acesso aos dados no storage e a opção de disco HOT-SPARE, onde o disco com problema é recriado. A **Figura 7-18** ilustra os componentes do storage inteligente.

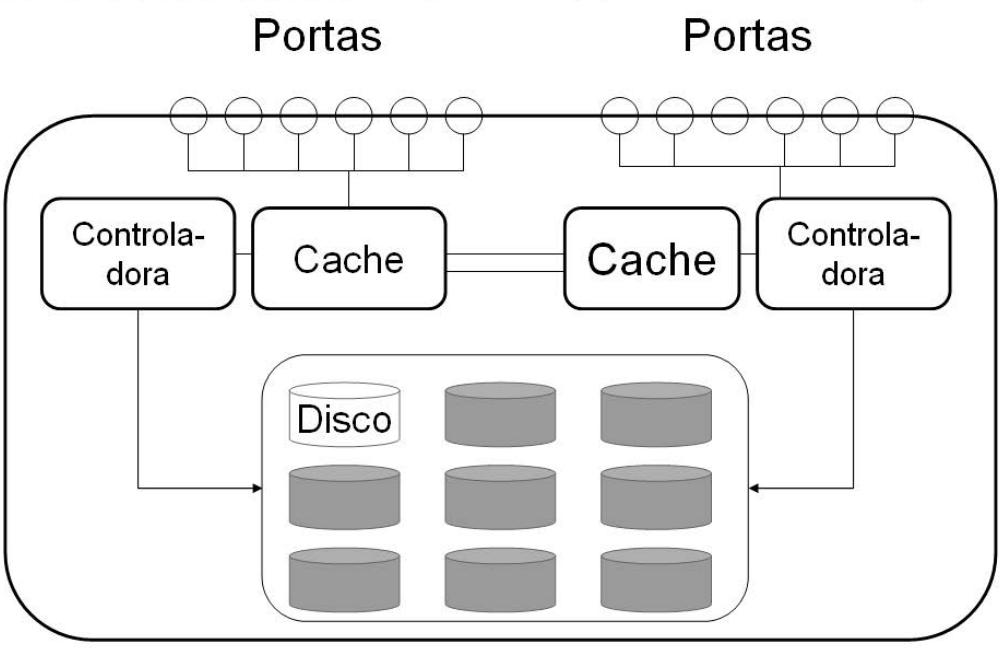


Figura 7-18 Storage Inteligente

7.6.7. Redes de Storage

O crescimento da massa de dado organizacional e as novas demandas relativas a regulamentação trouxeram à tona a necessidade da utilização de unidades de storage indepedentes do servidor . A maneira de conectar o storage evoluiu de soluções onde o storage era conectado diretamente ao servidor (*Direct Attached Storage* – DAS) para a criação de redes de storage indepedentes que usam uma combinação de protocolos específicos e interfaces de discos (*Storage Area Network* – SAN ou *Network Attached Storage* – NAS).

Quando se compara uma rede do tipo SAN a uma rede do tipo LAN é necessário entender algumas diferenças:

- O *overhead* de uma rede do tipo SAN é muito menor do que uma rede do tipo LAN.
- A proteção é uma característica muito mais necessária em uma rede LAN. O servidor acaba funcionando como um firewall para as redes SAN, o que simplifica o aspecto de segurança.
- Um comportamento mais inteligente para o controle do congestionamento na SAN é muito mais relevante do que na LAN, onde o TCP/IP simplesmente descarta o pacote, mas na SAN este aspecto acaba sendo um problema.

A **Figura 7-19** ilustra as principais técnicas de I/O utilizadas.

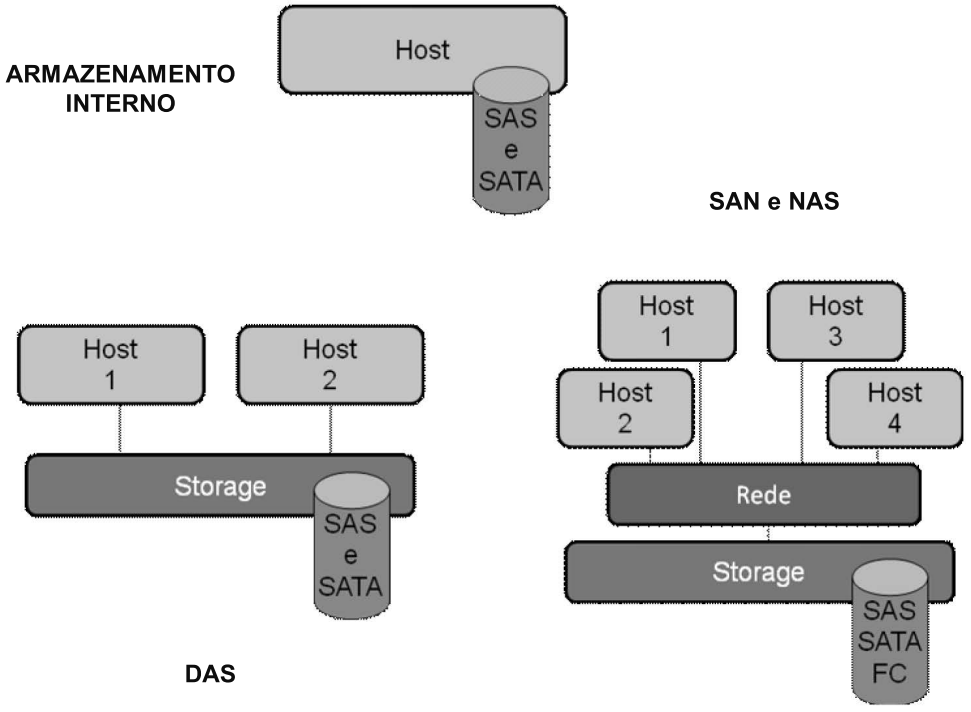


Figura 7-19 Técnicas de I/O

Nas redes SAN a infraestrutura de rede pode ser baseada no padrão FC ou Gigabit Ethernet. O armazenamento do tipo SAN (*storage area network*) é baseado em redes de armazenamento dedicadas e escaláveis que conectam servidores e dispositivos de storage usualmente no nível de bloco (dados de aplicação).

Nas redes NAS a infraestrutura é quase sempre baseada no padrão Gigabit Ethernet e o dado a ser armazenado é do tipo arquivo. O entendimento de quando utilizar uma infraestrutura ou outra é complexo e muitas vezes é confuso, causando problemas em projeto quando soluções, que deveriam ser baseadas em NAS, são baseadas em SAN e vice-versa. O padrão 10GbE também já é uma realidade.

O protocolo SCSI continua sendo o padrão utilizado na comunicação entre o servidor e o storage. Na prática, os protocolos FC e iSCSI encapsulam os comandos SCSI dentro do protocolo.

7.6.8. SAN FC

Os protocolos FC foram rapidamente adotados como uma arquitetura viável para aplicações que tratam de blocos em nível de I/O. O FC simplifica as ligações entre servidores e dispositivos de storage e diminui a perda de sinal, aumentando assim as distâncias máximas permitidas quando comparado ao SCSI convencional. Importante atentar que o FC é um protocolo que pode utilizar fibra óptica ou cobre como meio de comunicação.

O uso das redes SAN, baseadas no protocolo FC, possibilitou a ligação dos discos SCSI aos servidores, aumentando a velocidade e o número de dispositivos permitidos. Também adicionou suporte para protocolos de várias camadas de alto nível, incluindo SCSI, ATM e IP, sendo o SCSI o mais utilizado. Atualmente, com o surgimento de novas funcionalidades e dispositivos, as redes FC estão consolidadas. A **Figura 7-20** ilustra o protocolo FC.

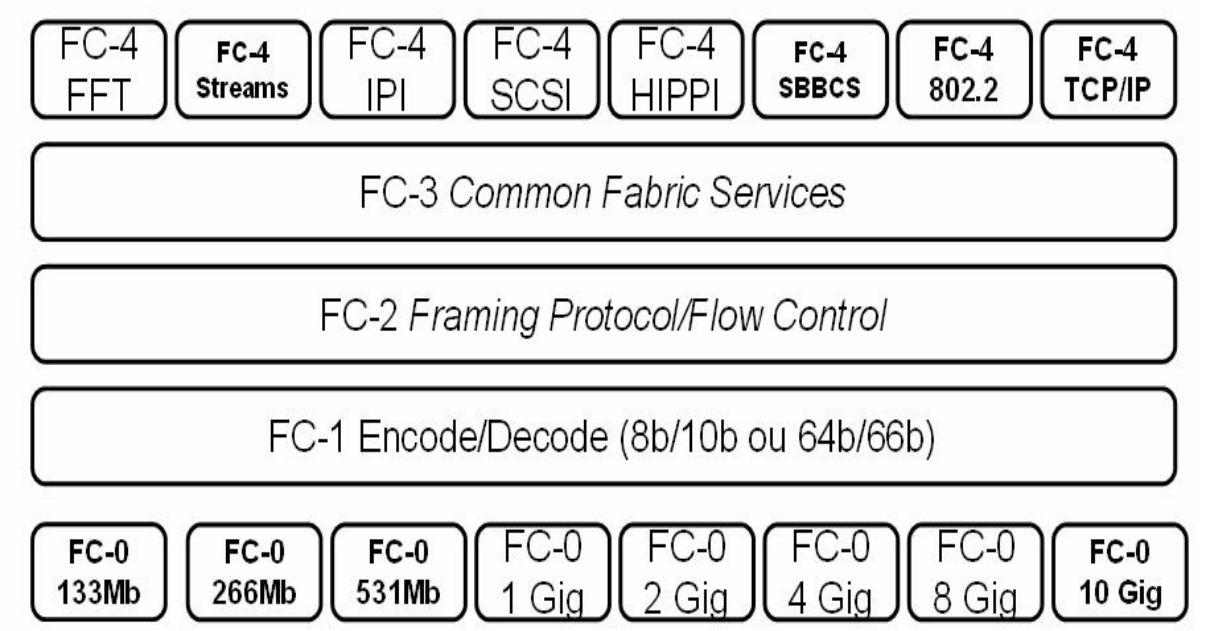


Figura 7-20 Protocolo Fibre Channel

As redes SAN FC possuem os seguintes componentes:

- **Portas dos nós (nodes ports):** nas redes FC os dispositivos são chamados de nós. Cada nó é fonte ou destino da informação para um ou mais nós. Cada nó requer uma ou mais portas para se comunicar com outros nós.

- **Cabeamento**: SANs utilizam cabos de fibra óptica ou de cobre para pequenas distâncias.
- **Dispositivos de interconexão**: os principais dispositivos de interconexão são os hubs, switches e directors. Hubs são pouco utilizados e compartilham a banda, devido aos dados serem transmitidos para todas as conexões. Switches são os componentes centrais de uma rede SAN e possuem função similar aos switches de LAN. Directors têm função similar aos switches, mas possuem um maior número de portas e maior robustez.
- **Unidades de armazenamento**: são normalmente unidades inteligentes com diversas funcionalidades já mencionadas anteriormente.
- **Software de gerenciamento da SAN**: gerencia a interface entre os servidores, dispositivos de interconexão e redes de storage.

Teoricamente uma rede SAN, baseada no protocolo Fibre Channel (FC), pode ter quinze milhões de nós interconectados. Os servidores, no caso do protocolo FC, utilizam placas do tipo HBA (*host bus adapter*) para permitir a conexão à SAN.

O protocolo FC é, de longe, o protocolo mais utilizado nas redes do tipo SAN e atende os requisitos de desempenho e confiabilidade necessários, sendo implementado em cinco camadas. As implementações iniciais ofereciam throughput de 100MB/s (1062.5 Mb/s), que já era bem superior ao padrão Ultra SCSI de 20MB/s usado em DAS até então.

O padrão FC pode operar em modo full-duplex chegando, neste caso, a 100MB/s. A última implementação de FC é de 8GB/s (1600MB/s ou 8.5Gb/s) comparada ao padrão atual Ultra320 SCSI, de 320MB/s.

7.6.9. SAN IP

Recentemente o padrão iSCSI (internet SCSI), que utiliza o padrão Ethernet para a comunicação na rede de storage, passou a ser uma opção.

O protocolo iSCSI é um meio de transportar os pacotes SCSI através do TCP/IP. Os principais componentes da rede iSCSI são servidores (*initiators*), *targets* e uma rede IP.

O protocolo iSCSI carrega os comandos SCSI do servidor (*initiator*) para os dispositivos *target* (storage). O iSCSI trabalha encapsulando os comandos SCSI dentro do TCP/IP e transportando-os através de uma rede IP. O tráfego iSCSI pode utilizar dispositivos Ethernet como switches e placas de rede (NICs) padrão para escoar o tráfego de dados em blocos como na SAN FC.

A **Figura 7-21** ilustra o protocolo iSCSI e sua relação com o TCP/IP.

Na comunicação SCSI existem as figuras do inicializador (*initiator*) e do destino (*target*). O *initiator* pode ser implementado por hardware no próprio adaptador, por software com TCP *offload* ou por software sem TCP *offload,* conforme ilustra a **Figura 7-22**.

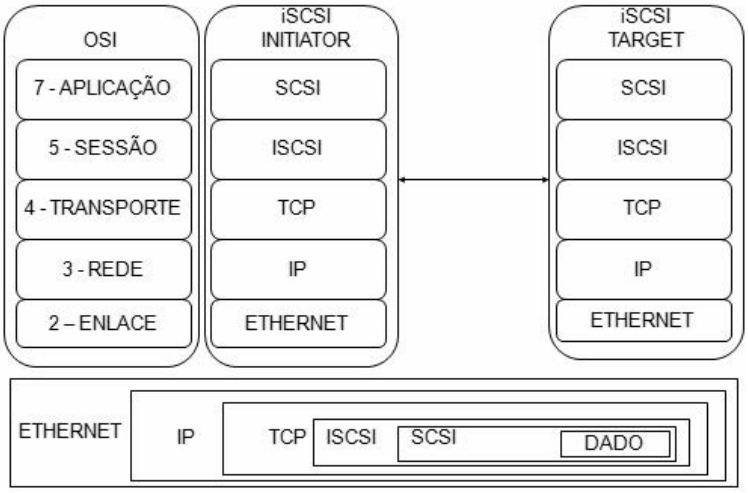


Figura 7-21 Protocolo iSCSI

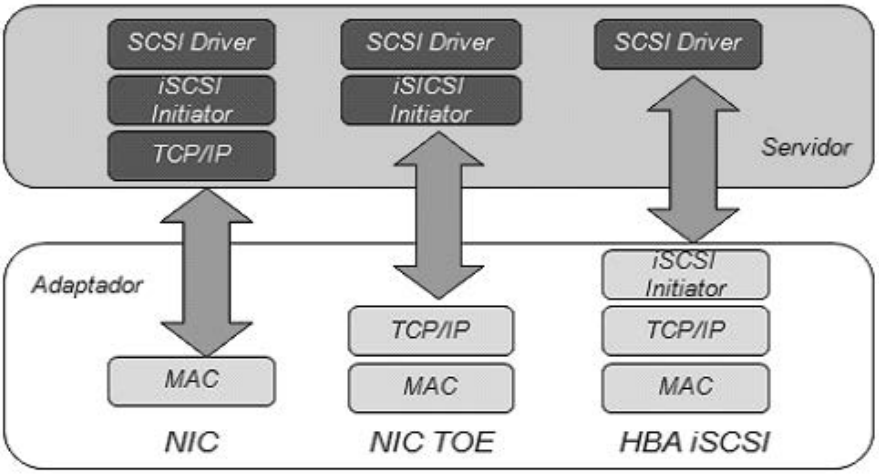


Figura 7-22 Implementação do iSCSI Initiator

Em geral o armazenamento em Internet SCSI (iSCSI) simplifica o armazenamento e reduz o custo de entrada em uma solução de storage e mesmo o TCO, quando comparado com o padrão FC. A adoção do padrão Ethernet pelo protocolo iSCSI permite um ganho de aprendizado em relação ao protocolo FC. Uma outra vantagem do iSCSI é que este pode ser incorporado a redes de armazenamento já existentes baseadas em NAS e em SAN FC, não obrigando uma ruptura tecnológica, desde que o storage já possua portas iSCSI, ou seja, realizado através da utilização de um *appliance*. Com a chegada do padrão 10 Gb/s Ethernet e a sinalização do aumento de banda para 40GB/s e 100Gb/s nos próximo anos, acredita-se que o padrão iSCSI tomará ainda mais espaço no mercado de redes de armazenamento. Importante ressaltar que a rede iSCSI, diferentemente da solução NAS, também trata de aplicações que utilizam blocos de dados como elementos padrão de I/O.

Os sistemas de arquivos utilizados por softwares de VIRTUALIZAÇÃO são construídos para dar suporte à VIRTUALIZAÇÃO. No caso do vSphere, tem-se as seguintes principais opções:

- **VMFS**: o VMware vSphere utiliza o VMFS em discos locais, volumes iSCSI ou volumes FC, criando um diretório para cada máquina virtual (VM). O VMFS é um sistema de arquivos clusterizado que pode ser usado simultaneamente por vários servidores ESX.
- **NFS**: o VMware vSphere pode utilizar um sistema de arquivos NFS contido em um servidor NFS. O vSphere monta o volume NFS criando um diretório para cada máquina virtual.

Uma decisão importante é a configuração do I/O de BLADES para soluções que utilizam o iSCSI e a VIRTUALIZAÇÃO. A **Tabela 7-3** sugere opções de I/O para configurações diferentes do BLADE DELL 1000E (ver **Figura 7-8**).

As configurações sugeridas são:

1. Configuração mínima de blade.
2. Maior banda para a LAN.
3. Configuração de LAN e SAN balanceada.
4. Maior banda para a rede iSCSI.
5. Fabrics isolados.

Tabela 7-3 iSCSI e LAN em blade 1000E

	1	2	3	4	5
MÓDULO A1	LAN	LAN	LAN	LAN	LAN
MÓDULO A2	LAN	LAN	LAN	LAN	LAN
MÓDULO B1	ISCSI SAN	ISCSI SAN	ISCSI SAN	ISCSI SAN	ISCSI SAN
MÓDULO B2	ISCSI SAN	ISCSI SAN	ISCSI SAN	ISCSI SAN	ISCSI SAN
MÓDULO C1	BLANK	LAN	LAN	ISCSI SAN	FABRIC ISOLADO
MÓDULO C2	BLANK	LAN	ISCSI SAN	ISCSI SAN	FABRIC ISOLADO

A decisão de utilizar uma rede SAN baseada no protocolo FC ou iSCSI com a camada de VIRTUALIZAÇÃO deve considerar a homologação dos produtos e as funcionalidades requeridas. Existe uma relação de interdependência entre o dispositivo de armazenamento e a virtualização, pois boa parte das funcionalidades obtidas com a VIRTUALIZAÇÃO depende da consolidação do armazenamento em um único dispositivo. As redes de storage do tipo SAN e NAS são suportadas pelos softwares de virtualização, o que permite otimizar a utilização de recursos.

A decisão entre utilizar hardware ou software *initiator* é essencialmente um *trade--off* entre preço e performance. A VMware mediu a diferença de performance utilizando o I/OMETER como gerador de WORKLOAD para medir a diferença entre as duas opções. A mesma WORKLOAD foi colocada em servidores idênticos utilizando recursos isolados de storage. O initiator por hardware disponibilizou 150% a mais de throughput e requereu só 25% dos recursos de CPU usados pelo software initiator. A diferença de preço entre um adaptador gigabit e um adaptador iSCSI ainda é alta.

7.6.10. Conceito na Prática: Dell Equallogic PS Series

A linha Dell Equallogic PS é baseada em iSCSI e totalmente compatível com os principais softwares de VIRTUALIZAÇÃO do mercado. A ideia central é permitir que empresas tenham uma boa opção de storage para redes SAN virtualizadas sem grandes custos de administração e também sem perder os requisitos essenciais de desempenho e escalabilidade. A capacidade do array Equallogic vai de 400 GB até 48 TB em chassis de 3U e 4U.

A linha Dell Equallogic é modular e permite que o gerente de TI adquira só o storage necessário e que depois possa expandi-lo fazendo o balanceamento de carga através de todos os recursos disponíveis na rede SAN iSCSI montada com as unidades Equallogic. Cada array possui tolerância a falhas e componentes *hot-swappable.*

Um diferencial da linha Equallogic é que as funcionalidades de software já estão todas disponíveis sem ter que pagar licença adicional. O array Equallogic permite o crescimento modular *on-demand,* tornando o ambiente de infraestrutura de TI mais flexível. Cada storage PS utiliza redes Ethernet padrão e o protocolo iSCSI. Os switches da rede iSCSi são convencionais e os discos podem ser Serial ATA (SATA), Serial Attached SCSI (SAS) e Solid State Disk (SSD). Os servidores podem utilizar placas de rede convencional para acessar a rede iSCSI ou mesmo uma HBA iSCSI. A série PS pode ser utilizada também em DAS e NAS, além da opção iSCSI suportando um fabric em 10GbE.

O Equallogic trabalha com o conceito de grupo de arrays. O grupo é visto com um único array pelo administrador e pelos servidores da rede. Quando um array é configurado como um membro do grupo, seus discos em RAID são adicionados ao pool de armazenamento. O grupo aparece na rede para os servidores clientes como uma única entidade que oferece acesso à rede de storage no modo bloco – tipicamente utilizado em redes SAN. Cada membro do grupo no conceito do Equallogic trabalha cooperando com os outros membros e otimizando o uso dos recursos de armazenamento, independente da VIRTUALIZAÇÃO. Os volumes são criados nos espaços disponíveis dos pools de storage arrays. Snapshots e réplicas do volume podem ser criadas também. Os espaços em disco criados são particionados em unidades chamadas de páginas. As páginas são movimentadas di-

namicamente entre os membros do grupo. Vale ressaltar que nas operações de expansão dos grupos os dados permanecem disponíveis.

Um aspecto interessante do Equallogic é que tradicionais operações de SAN como criação de clones e snapshots, que dependem de um software específico, já estão incluídas no produto sem custo adicional. Os snapshots permitem realizar backups online e a capacidade de replicação do storage permite criar um site para recuperação a desastres a custo interessante, quando comparado a outras soluções de mercado.

As principais funcionalidades da série PS podem ser resumidas abaixo:

- Componentes redundantes no array incluem discos, controladoras, interfaces de rede, fontes de energia e ventiladores.
- Hardware *hot swappable*.
- Perdas das controladoras não acarretam perda da proteção de RAID.
- Discos são automaticamente configurados no nível de RAID adequado e hot spares são reservados.
- Hot sparing e recuperação de dados acontecem sem intervenção humana.
- Replicação do volume propicia proteção de desastres no nível do site.
- Upgrade do hardware de um membro do grupo pode ser feito enquanto o grupo está online simplesmente removendo o array do grupo. Um membro pode ser retirado do grupo sem interrrupção.
- Recursos de Thin Provisioning já fazem parte da oferta.

A **Figura 7-23** ilustra o Dell Equallogic 6010.

Figura 7-23 Dell Equallogic 6010

7.6.11. NAS

O NAS (*network attached storage*) é baseado em redes de storage IP e é primariamente utilizado para compartilhamento de arquivos. Quando comparado ao DAS, é mais escalável e possui melhor disponibilidade, além de ser mais fácil de gerenciar.

Normalmente o seu uso e gerenciamento requerem maior investimento inicial e conhecimento mais especializado.

O NAS usa protocolos de rede e de compartilhamento de arquivos. Estes protocolos incluem o TCP/IP para transferência de dados e dos protocolos CIFs e NFS para serviços de arquivo remoto. Usuários do Windows e do UNIX podem compartilhar os mesmos dados armazenados em um servidor NAS. NAS são acessados por clientes e servidores em uma rede IP. Muitas vezes o servidor NAS utiliza múltiplas interfaces de rede.

O NAS utiliza seu próprio sistema operacional normalmente dedicado ao trabalho de servir arquivos, ou seja, para otimizar o I/O, para vários sistemas operacionais. Logicamente o NAS pode servir mais clientes do que um servidor de arquivos convencional.

Existe uma relação de interdependência entre o dispositivo de armazenamento e a VIRTUALIZAÇÃO. Boa parte das funcionalidades obtidas com a VIRTUALIZAÇÃO depende da consolidação do armazenamento em um único dispositivo. As redes de storage do tipo SAN e NAS são suportadas pelos softwares de VIRTUALIZAÇÃO, o que permite otimizar a utilização de recursos. Também diversas funcionalidades oferecidas dependem da plena interoperabilidade entre os dispositivos de storage e os softwares de VIRTUALIZAÇÃO.

7.6.12. SAN FCoE

Os padrões e tecnologias de rede SAN vêm sendo alterados ao longo dos últimos quinze anos. A utilização de um único padrão é um sonho. Na prática, as organizações já possuem redes de armazenamento e padrões estabelecidos, e novos padrões estão chegando. O segredo do FCoE é integrar novas opções com arranjos já instalados e maduros, evitando o rompimento tecnológico precoce. Em resumo:

- Fibre Channel (FC) é protocolo de armazenamento de bloco padrão para redes de storage (SAN). Em DATACENTERS corporativos, aproximadamente 90% dos dispositivos de armazenamento são conectados via FC. A SAN permite consolidação do storage, gerenciamento centralizado, alta performance e confiabilidade e reconfiguração rápida. Mesmo assim FC não conseguiu ser o padrão para fabrics de clusters e redes de volume, devido à vantagem do padrão Ethernet em termos de custo. Mesmo indo para 8Gb/s, o padrão FC será mais lento do que o padrão Ethernet, que deve se mover rapidamente para 40Gb/s e 100Gb/s nos próximos anos.
- iSCSI transporta SCSI over TCP/IP, possibilita que SANs sejam criadas utilizando switches Ethernet de baixo custo e softwares iSCSI initiators. Tem sido utilizado como um protocolo para blocos em redes SMB em 1Gb/s e agora também com adaptadores Ethernet em 10Gb/s, utilizando TOE (TCP offload engines).
- SAS (*serial attached SCSI*) é uma alternativa focada em custo que permite trocar o SCSI paralelo. Mas apresenta limites de escalabilidade, performance e distância, quando comparada com FC em redes de storage.
- NAS appliances utilizam protocolos NFS e CIFS para sistemas de arquivos acessados através de redes Ethernet.

Por que então utilizar mais um padrão? A ideia do padrão FCoE é transportar os frames FC através de um fabric Ethernet de 10Gb/s, possibilitando a melhoria da performance da aplicação enquanto reduz custo, energia e gerenciamento, convergindo storage, rede e clusters em um único *fabric*.

Importante reforçar que iSCSI e FCoE são padrões distintos. FCoE é uma alternativa ao iSCSI.

A **Figura 7-24** ilustra o mapeamento FCoE.

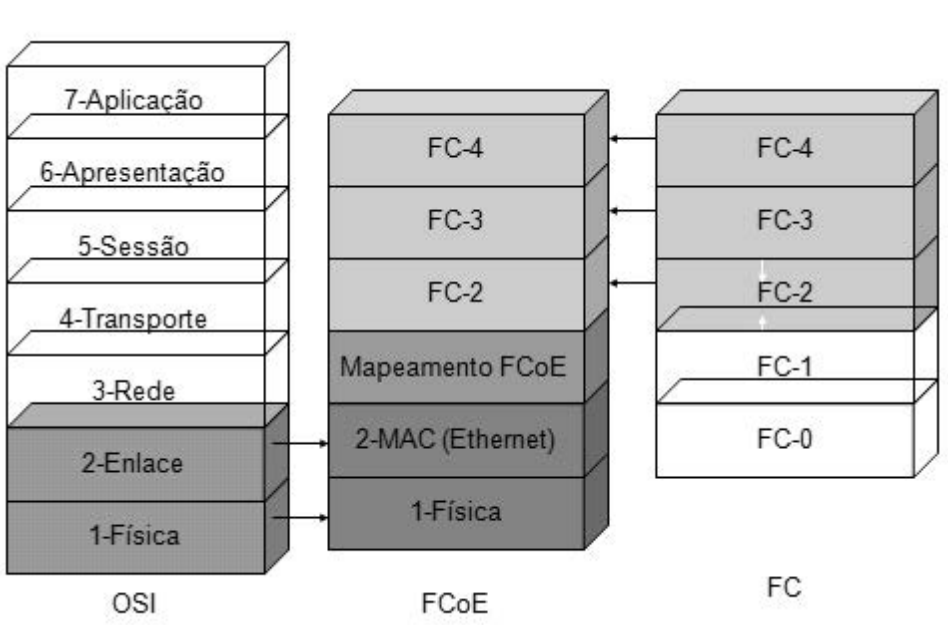


Figura 7-24 Mapeamento Fibre Channel para Ethernet

Os novos switches baseados em FCoE trazem o conceito inovador de *unified fabric*, conjunto de tecnologias que permitem a consolidação das diversas redes existentes em um DATACENTER com uma única infraestrutura de rede.

Com a tecnologia *unified fabric*, as diversas redes Ethernet (LAN), redes de armazenamento Fibre Channel (SAN) e redes de computação de alto desempenho (HPC) podem ser implementadas sobre uma infraestrutura Ethernet unificada, permitindo maior flexibilidade, compartilhamento de recursos e redução de custos com cabeamento, energia e equipamentos dedicados.

No caso de blades a situação é ainda mais crítica, pois a exigência de mais portas de I/O para maior número de cores por processador e maior uso de máquinas virtuais reforça a necessidade de integração de tecnologias.

O *unified fabric* reduz o custo eliminando a necessidade de utilizar múltiplos adaptadores para dados específicos como FC para storage e Infiniband para clustering. Enquanto o padrão 10Gbps permite obter alta performance com o FCoE, o FCoE também pode rodar em redes existentes FC sem custos adicionais. O projeto Open-FC.org está criando os drivers para o SO Linux. A **Figura 7-25** ilustra a rede FcoE.

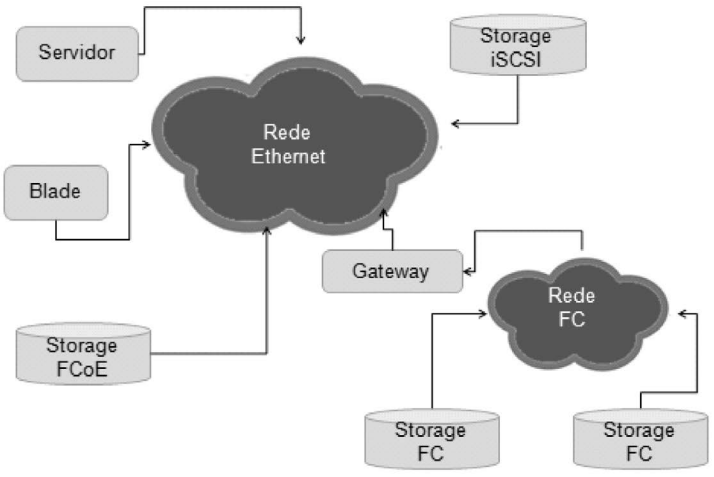


Figura 7-25 SAN FCoE

Os switches unificados adotam o padrão DATACENTER Ethernet com melhorias no controle de fluxo e gerenciamento do congestionamento da rede (padrão 802.1p QoS). Os switches encapsulam o frame FC no pacote Ethernet (FCoE), suportando a consolidação do I/O.

O padrão FCoE permite conduzir o tráfego Fibre Channel através de uma rede Ethernet. A **Figura 7-26** ilustra a ideia de redução dos adaptadroes no padrão FCoE. Substitui-se a LAN e a HBA por um cartão CNA (adaptador de rede convergente- converged network adapter).

A unificação das redes SAN e LAN decorre das tecnologias Ethernet e FCoE (Fibre Channel over Ethernet), onde a rede Ethernet passa a ser o novo meio físico das redes lógicas LAN e SAN. Os servidores possuem acesso à LAN e à SAN pelo adaptador CNA, eliminando a necessidade de NIC para LAN e HBA ou outra NIC para SAN.

Os resultados da convergência LAN/SAN são: custos menores com infraestrutura de rede, cabeamento, energia, resfriamento, gerenciamento, além da consolidação em si.

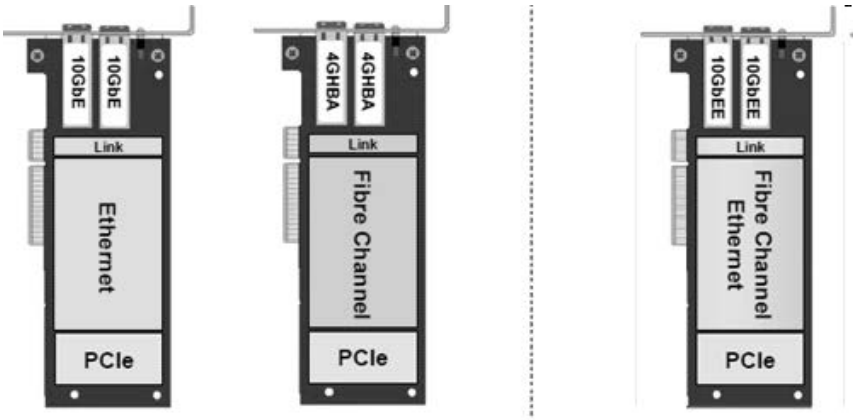


Figura 7-26 NIC, HBA e CNA

7.6.13. Conceito na Prática: Unified Storage EMC Celerra NS-120

Os sistemas EMC Celerra NS-120 podem ser os elementos de integração de uma estratégia abrangente de gerenciamento do ciclo de vida das informações – estratégia que ajuda a empresa a obter o maior valor possível de suas informações, com o menor TCO (*Total Cost of Ownership*, custo total de propriedade), em cada fase do ciclo de vida das informações. O gerenciamento do ciclo de vida das informações associa o nível de serviço ao aplicativo, pelo custo adequado – no momento certo.

Entre os recursos do CELERRA destacam-se:

- Snapshots e gerenciamento de snapshots para criação de visualizações lógicas somente leitura e de leitura/gravação para a proteção de dados e testes de aplicativos.
- Gerenciamento automático de volumes para habilitar o provisionamento baseado em perfis.
- Provisionamento virtual (thin) para aumentar a utilização da capacidade.
- Deduplicação de dados para aumentar a eficiência do armazenamento.
- APIS para movimentação de arquivos para fins de arquivamento e de armazenamento classificado por níveis.

A **Figura 7-27** ilustra o EMC CELERRA e as opções de conectividade (a opção FCoE ainda não tinha sido lançada pelo fabricante).

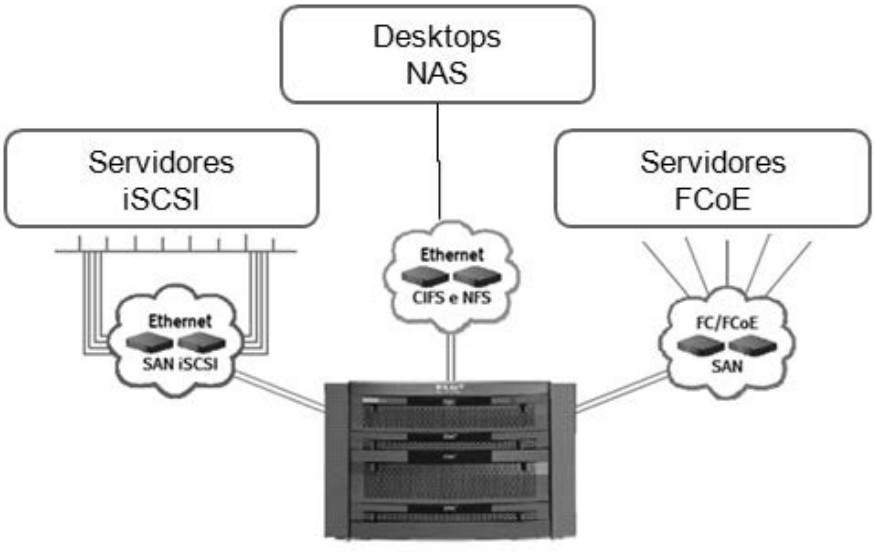


Figura 7-27 Opções de Conectividade no EMC CELERRA NS-120

Cada sistema CELERRA consiste em um ou mais X-Blades. X-Blades são componentes do gabinete que executam seu próprio SO e que recuperam dados de um dispositivo de armazenamento, disponibilizando-os aos clientes da rede. Os X-Blades fornecem à interface a rede IP e são responsáveis por apresentar os compartilhamentos de NAS ou

LUNs iSCSI. A **Figura 7-28** ilustra os componentes de uma arquitetura CELERRA. O DME (*Data Mover Enclosure*) é um gabinete que contém um ou dois X-Blades *(Data Movers*), módulos de gerenciamento e módulos de alimentação e refrigeração.

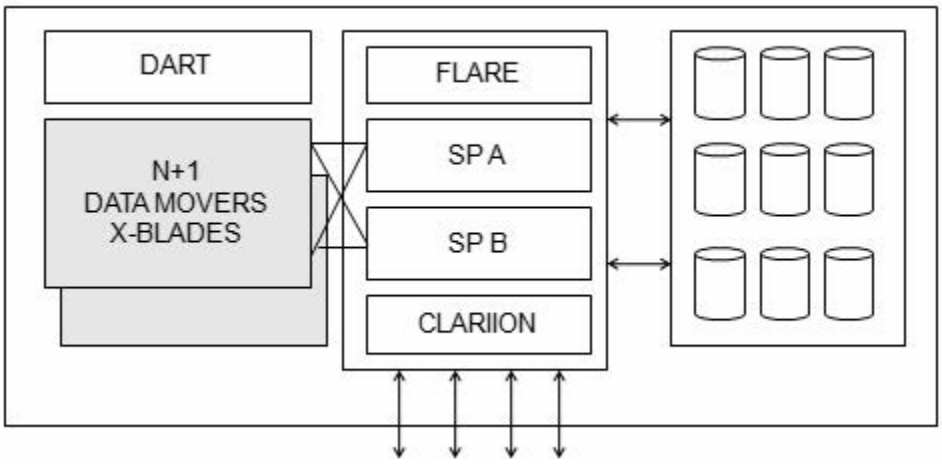


Figura 7-28 Componentes da Arquitetura CELERRA

A plataforma de armazenamento unificado NS-120 tem suporte para as configurações de X-Blade simples e dupla. As configurações de X-Blades duplas podem ser implantadas no modo principal/principal para oferecer o mais alto desempenho ou no modo principal/standby como medida adicional de proteção para disponibilidade de hardware (ou seja, failover de X-Blade). O NS-120 usa a tecnologia de array de back-end do EMC CLARiiON CX4.

Cada X-Blade consiste em:

- Dois processadores Intel Xeon LV de 2,8 GHz.
- RAM Double Data Rate de 4 GB (266 MHz) em um FSB de 800 MHz.
- Duas portas ópticas Fibre Channel de 4 Gb/s para conectividade com switch.
- Até duas portas Fibre Channel de 4 Gb/s para conectividade de fita.
- Quatro portas 10/100/1000 BaseT ou duas portas ópticas 10 Gigabit Ethernet e duas portas 10/100/1000 BaseT.
- Uma porta de gerenciamento 10/100/1000.
- Instância do software DART File Server.

Os X-Blades executam o DART (*Data Access in Real Time*), um sistema operacional criado especificamente para o gerenciamento de armazenamento. O DART é independente do hardware e do CELERRA.

O SPE (*Storage Processor Enclosure*), que é o gabinete principal do sistema CLARiiON, utiliza duas controladoras de armazenamento (*Storage Processors* – SPs) ativas para I/O de disco. É conectado a um par redundante de barramentos Fibre Channel de 4 Gb/s, fornecendo acesso contínuo do drive aos hosts em caso de falha de uma con-

troladora de armazenamento ou de um barramento. O NS-120 requer, no mínimo, cinco drives (Fibre Channel ou SATA de 7.200 rpm) e aceita, no máximo, 120 drives de disco em oito chassis para expansão de disco. O SPE executa o FLARE e possui uma implementação robusta de RAID. Os compartimentos de disk array são gabinetes de 15 discos (FC, SATA, Flash). Os DPEs (*Disk Processor Enclosure*) permitem a expansão de discos no sistema CLARiiON.

A tecnologia ULTRAFLEX oferece um sistema de armazenamento que pode ser personalizado com facilidade preenchendo os slots de I/O com módulos de I/O que atendam às necessidades de conectividade, dinâmicas e em expansão. A **Figura 7-29** ilustra o acréscimo do módulo de I/O ao CELERRA.

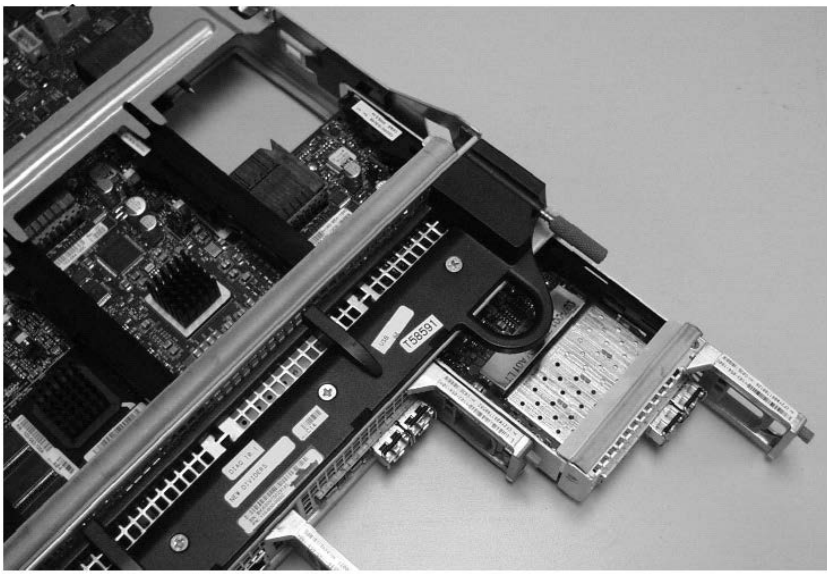


Figura 7-29 Acréscimo do Módulo de I/O ao CELERRA

A interface do usuário para o CELERRA é a Control Station, que possui conectividade SSH e GUI. A Control Station é baseada em uma CPU Celeron de 2 GHz.

A nova versão do EMC CELERRA a ser lançada ainda este ano pela EMC possuirá as seguintes funcionalidades:

- **Unisphere Manager:** ponto de controle simples e unificado para gerenciamento de dados em acesso a bloco ou arquivo.
- **Sub-Lun FAST:** movimentação de dados entre camadas do storage baseadas em políticas e acesso a disco(I/O).
- **Integração nativa com o VMware:** visualizar, provisionar e gerenciar serviços de armazenamento via vSphere.
- **vSphere VAAI ready:** integração com o VMware para acelerar operações de armazenamento.
- **Compactação de LUNs:** recurso de eficiência no uso de espaço para dados em bloco ou arquivo.

7.6.14. VIRTUALIZAÇÃO do Storage

A VIRTUALIZAÇÃO do *storage* é a abstração lógica do sistema físico do *storage*. Diferentemente da virtualização dos servidores, facilmente justificável com uma boa análise de TCO, conforme será visto no próximo capítulo. A VIRTUALIZAÇÃO do *storage* não é tão popular quanto a VIRTUALIZAÇÃO de servidores, pois faz mais sentido para organizações que possuem sistemas de *storage* heterogêneos, o que, em geral, só é comum em grandes empresas.

Para organizações que possuem sistemas heterogêneos, a VIRTUALIZAÇÃO do *storage* permite simplificar a administração e reduzir o custo de gerenciamento do sistema de *storage* como um todo. A virtualização do *storage* também facilita obter ou mesmo melhorar o nível de interoperabilidade entre unidades de *storage* independentes.

A ideia da VIRTUALIZAÇÃO do *storage* é simples. O sistema de *storage* passaria a funcionar como um sistema único do ponto de vista do administrador. Os recursos disponíveis então poderiam ser mais bem aproveitados independentes da plataforma de *storage*.

Tipos de Virtualização do Storage

A VIRTUALIZAÇÃO trata de providenciar um *storage* lógico para servidores e aplicações independentes dos recursos físicos. A VIRTUALIZAÇÃO do *storage* pode acontecer em um ambiente de SAN no nível de bloco ou em um ambiente de NAS no nível de arquivo.

- Na VIRTUALIZAÇÃO no nível de bloco, o storage virtual se apresenta para os servidores na forma de discos virtuais.
- Na VIRTUALIZAÇÃO no nível de arquivo, o storage virtual entrega para os servidores arquivos e diretórios.

O conceito de VIRTUALIZAÇÃO do *storage* pode incluir discos, unidades de fita, sistemas de arquivos e até mesmo os arquivos. A virtualização pode ocorrer no servidor, no *storage*, ou na rede via switches inteligentes ou appliances. A VIRTUALIZAÇÃO pode ocorrer in-band (dados e controle no mesmo percurso) ou out-of-band (dados e controle em percursos distintos).

A VIRTUALIZAÇÃO do *storage* requer automação das rotinas e deve também ser baseada em políticas para reduzir a intervenção manual. Existem também algumas iniciativas para tornar as aplicações prontas para a VIRTUALIZAÇÃO que propiciam melhor interface entre o SO e os utilitários de *storage*.

Appliances de Virtualização

Appliance é um equipamento desenvolvido e configurado para executar uma função específica dentro de um sistema.

A VIRTUALIZAÇÃO de storage utilizando *appliances* é a mais comum, pois propicia independência do servidor e do *storage*. Os appliances existentes são baseados em hardware ou em software rodando em servidor *wintel* e propiciam interconexão

com diferentes sistemas operacionais, plataformas de servidores e *storage* de diferentes fornecedores.

Os *appliances* disponíveis podem suportar protocolos FC e iSCSI para acesso dos servidores e *storage*s baseados em discos SCSI,SAS,SATA, FC.

A utilização de *appliances* iSCSI pode repor sistemas antigos baseados em DAS em uma solução mais econômica e propiciar um caminho de migração de pequenas redes para grandes redes de *storage*. Os appliances podem dar suporte à VIRTUALIZAÇÃO in-band e out-of-band.

A VIRTUALIZAÇÃO *in-band* pode dividir a rede de *storage* em dois *fabrics* separados: um para a conectividade do servidor ao *appliance* e outro para a conectividade do *appliance* ao *storage*. O dispositivo deve residir entre os servidores e o *storage*. Em ambientes FC serão necessárias duas *fabrics* para conectividade dos servidores e outra para os *storage*s. O *appliance* neste tipo de configuração é *target* para os servidores e *initiator* para os dispositivos de *storage*.

Em algumas situações o *appliance* pode se tornar um gargalo do desempenho nesta modalidade de virtualização e também pode apresentar latência para o transporte dos dados que agora tem que passar por duas redes.

Uma vantagem deste tipo de VIRTUALIZAÇÃO é o fato de propiciar separação física entre os servidores e o *storage*, tornando a rede mais segura e dificultando o acesso não autorizado de servidores aos *storages*.

A VIRTUALIZAÇÃO *out-of-band* usa percursos separados para o controle e os dados. Neste caso, o appliance é uma espécie de peer da rede. Neste caso, os servidores acessam os dados diretamente através da rede SAN. O possível gargalo existente nas redes in-band é evitado neste arranjo e o desempenho depende do *storage* e dos switches da SAN.

A VIRTUALIZAÇÃO *out-of-band* requer a instalação de drivers nos servidores para manter o mapeamento dos blocos que fazem parte da virtualização e exceções gerados pelos *appliances.* A grande vantagem das soluções out-of-band é a que elas não requerem mudança na infraestrutura de SAN já existente, facilitando o processo de migração para redes de *storage* virtualizadas.

A **Figura 7-30** indica as duas possíveis utilizações. As linhas tracejadas são referentes à conexão out-of-band.

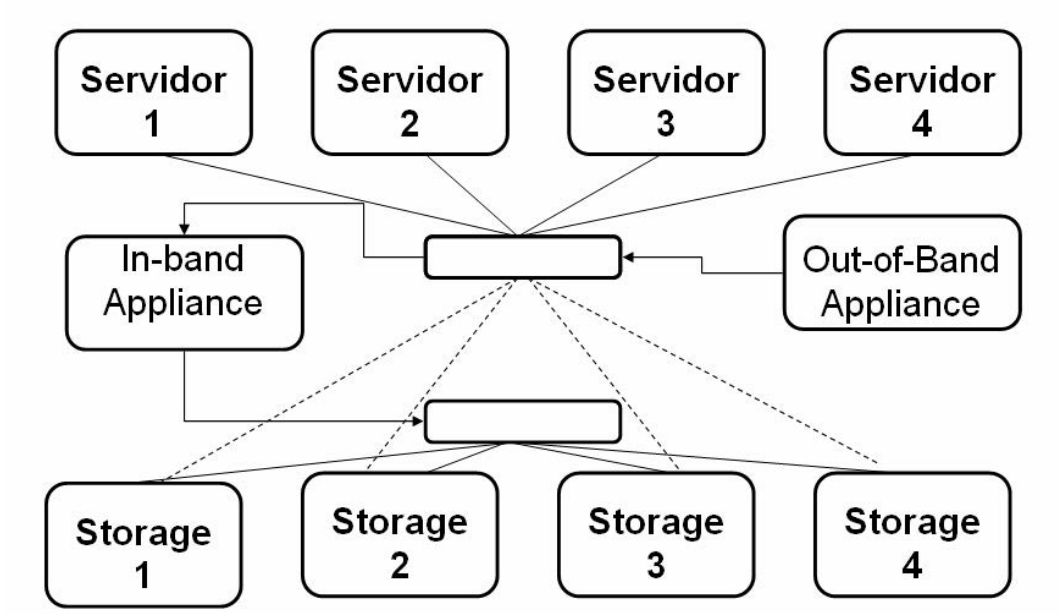


Figura 7-30 Appliances In-Band ou Out-of-Band de VIRTUALIZAÇÃO do Storage

Como o appliance acaba sendo um ponto de falha, é preciso pensar nestes casos em um arranjo de alta disponibilidade (HA).

- No caso da VIRTUALIZAÇÃO in-band, a alta disponibilidade implica em duplicar switches, appliances e paths.
- No caso da virtualização out-of-band, é menos complicado, pois os appliances precisam só serem conectados a uma rede SAN dual-pathed.

Ambos os tipos de rede podem conter tecnologias de alta disponibilidade como servidores em cluster, replicação e snapshots.

7.6.15. Conceito na Prática: EMC VPLEX

A EMC lançou recentemente a família VPLEX com foco na mobilidade das informações e acesso a elas nos DATACENTERS. A plataforma oferece agrupamento local e distribuído.

- O agrupamento local proporciona cooperação e mobilidade imperceptíveis de elementos físicos em um local.
- O agrupamento distribuído amplia o acesso entre dois locais à distância. O VPLEX é uma solução para grupar armazenamento EMC e não EMC.

O AcessAnywhere disponível com o VPLEX é a tecnologia que permite o compartilhamento, acesso e remanejamento de uma só cópia de dados à distância. O EMC *GeoSyncrony* é o sistema operacional do VPLEX.

Assim como o VPLEX remove barreiras físicas e possibilita que os usuários acessem uma única cópia dos dados em diferentes locais geográficos, ele também habilita

clusters físicos e virtuais de hosts estendidos geograficamente. Isso permite um compartilhamento imperceptível de carga entre vários locais, oferecendo a flexibilidade de remanejar cargas de trabalho entre locais antes de eventos planejados. Além disso, no caso de um evento não planejado que poderia provocar interrupção em um dos DATACENTERS, os serviços com falha podem ser reiniciados no local sobrevivente com o mínimo de esforço e, ao mesmo tempo, reduzindo o tempo de recuperação.

O EMC VPLEX muda completamente o modo de gerenciar e fornecer TI – particularmente quando implantado com VIRTUALIZAÇÃO de servidor. Habilitando novos modelos de computação para a operação e o gerenciamento de TI, os recursos podem ser agrupados – em pool e em colaboração em toda a pilha – com a capacidade de mover dinamicamente os aplicativos e dados entre as regiões e provedores de serviço.

A **Figura 7-31** ilustra a ideia do VPLEX.

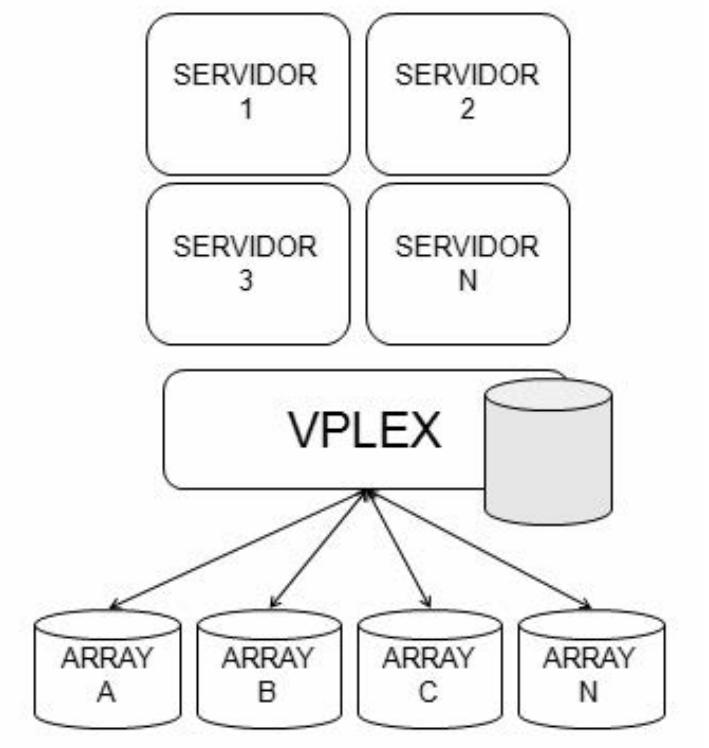


Figura 7-31 EMC VPLEX

Criado a partir de mecanismos de processador redimensionáveis e disponíveis, o EMC VPLEX foi desenvolvido para permitir o redimensionamento de pequenas para grandes configurações. O VPLEX reside entre os servidores e os ativos de armazenamento heterogêneo e usa uma arquitetura única de cluster que permite que os servidores dos vários DATACENTERS tenham acesso de leitura/gravação a dispositivos compartilhados de armazenamento em bloco. As características únicas desta nova arquitetura abrangem:

- Hardware de clustering scale-out, que possibilita que você comece pequeno e cresça com níveis de serviço previsíveis.

- Cache avançado de dados usa cache SDRAM de grande escala para aumentar o desempenho e reduzir a latência de I/O e o conflito de acesso de arrays.
- Coerência de cache distribuído para compartilhamento, balanceamento e failover automáticos de I/O no cluster.
- Visão consistente de um ou mais LUNs nos clusters VPLEX – tanto separados por alguns centímetros em um DATACENTER quanto entre distâncias síncronas – possibilitando novos modelos de alta disponibilidade e remanejamento de cargas de trabalho.

Com uma arquitetura *scale-up* e *scale-out*, o cache avançado de dados e a coerência de cache distribuído do VPLEX oferecem capacidade de recuperação de cargas de trabalho e também compartilhamento, balanceamento e failover automáticos de domínios de armazenamento, além de permitir o acesso local e remoto aos dados com níveis previsíveis de serviço.

A família VPLEX é composta por dois produtos: EMC VPLEX Local e EMC VPLEX Metro.

- O EMC VPLEX Local oferece agrupamento local, com gerenciamento simplificado e mobilidade de dados que não causa interrupções entre arrays heterogêneos.
- O EMC VPLEX Metro oferece agrupamento distribuído, que proporciona mobilidade de dados e acesso a eles entre dois clusters VPLEX em distâncias síncronas.

Uma configuração de VPLEX Local é definida por até quatro mecanismos VPLEX, que são integrados em uma única imagem de cluster por meio de interconexões de malha entre mecanismos totalmente redundantes. Esse recurso de interconexão de cluster permite o acréscimo online de mecanismos VPLEX, oferecendo capacidade excepcional de expansão às configurações do VPLEX Local e do VPLEX Metro. Toda a conectividade entre os nós de cluster VPLEX e entre as configurações do VPLEX Metro é totalmente redundante, garantindo proteção contra pontos únicos de falha.

Pode-se fazer o scale-up de um cluster VPLEX por meio do acréscimo de mais mecanismos e o scale-out pela conexão de clusters em um Metro-Plex (dois clusters VPLEX Metro conectados em distâncias longas). O VPLEX Metro ajuda a transferir e compartilhar cargas de trabalho – inclusive entre hosts virtualizados –, consolidar DATACENTERS e otimizar a utilização de recursos entre DATACENTERS de modo imperceptível. Além disso, ele oferece mobilidade de dados, gerenciamento de armazenamento heterogêneo e maior disponibilidade dos aplicativos sem causar interrupções. O VPLEX Metro tem suporte para até dois clusters, que podem estar no mesmo DATACENTER em dois locais diferentes em distâncias síncronas (aproximadamente até 60 milhas ou 100 quilômetros).

Em um único local ou entre locais, a família VPLEX aumenta a capacidade de recuperar dados e armazenamento. O VPLEX permite que se espelhem volumes nos locais e entre eles, fornecendo disponibilidade contínua dos aplicativos em caso de uma falha de componente. Esse recurso pode aumentar a proteção e a disponibilidade de aplicativos essenciais, além de aproveitar os recursos de armazenamento atuais – sem exigir recursos de host.

O VPLEX Metro, combinado com o VMware e o Distance VMotion, fornece um recurso único com o qual pode-se fazer a transferência e o remanejamento, de modo imperceptível, de máquinas virtuais e seus respectivos aplicativos e dados, à distância. A EMC Global Solutions desenvolveu uma arquitetura de referência que demonstra uma configuração do VPLEX Metro combinada com VMware e Distance VMotion para mobilidade de aplicativos sem causar interrupções locais, oferecendo a flexibilidade de remanejar cargas de trabalho entre locais antes de eventos planejados. Além disso, no caso de um evento não planejado que poderia provocar interrupção em um dos DATA-CENTERS, os serviços com falha podem ser reiniciados no local sobrevivente com o mínimo de esforço e, ao mesmo tempo, reduzindo o tempo de recuperação.

O VPLEX ajuda a mudar o modo de gerenciar e fornecer TI – particularmente quando implantado com a VIRTUALIZAÇÃO. Habilitando novos modelos de computação para a operação e o gerenciamento de TI, os recursos podem ser agrupados – em pool e em colaboração em toda a pilha – com a capacidade de mover dinamicamente os aplicativos e dados entre as regiões e provedores de serviço.

O VPLEX evoluirá para uma solução assíncrona (VPLEX Geo) e para uma solução anywhere (VPLEX Global).

7.7. Alta Disponibilidade e Recuperação de Desastres e VIRTUALIZAÇÃO

A alta disponibilidade (HA) e a recuperação de desastres (DR) devem ser pensados como um continuum e acontecem em várias camadas, onde cada camada propicia os níveis de disponibilidade adequados para a camada superior. As camadas são descritas a seguir e ilustradas na **Figura 7-32**.

- **Plataforma (dentro da caixa)**: servidores com redundância de fonte e discos, por exemplo.
- **Dado**: redes SAN redundantes, por exemplo.
- **Aplicação**: cluster failover de aplicação, por exemplo.
- **Site**: replicação de site, por exemplo.

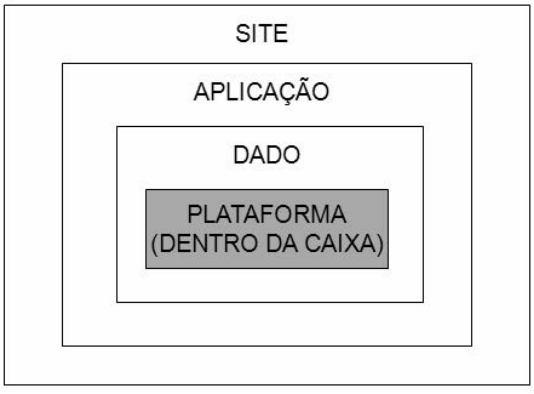


Figura 7-32 Como pensar a Alta Disponibilidade e a Recuperação de Desastres

Ou seja, não se deve pensar a alta disponibilidade e a recuperação de desastres em um único nível. A ideia de fazer um cluster para a aplicação, por exemplo, não deve eliminar a redundância nas fontes de alimentação e o nível de RAID utilizados em um conjunto de disco dentro dos servidores. Da mesma forma, a replicação entre sites não deve eliminar a necessidade de se fazer o cluster da aplicação no local utilizando técnicas de clusters de servidores.

Uma forma hierárquica de relacionar a continuidade do negócio (*business continuity*) com a recuperação de desastres e a alta disponibilidade da TI é mostrada na **Figura 7-33**. A figura reforça a ideia de considerar a alta disponibilidade, backup & restore e a replicação como um subconjunto da recuperação, que, por sua vez, é um subconjunto da continuidade do negócio.

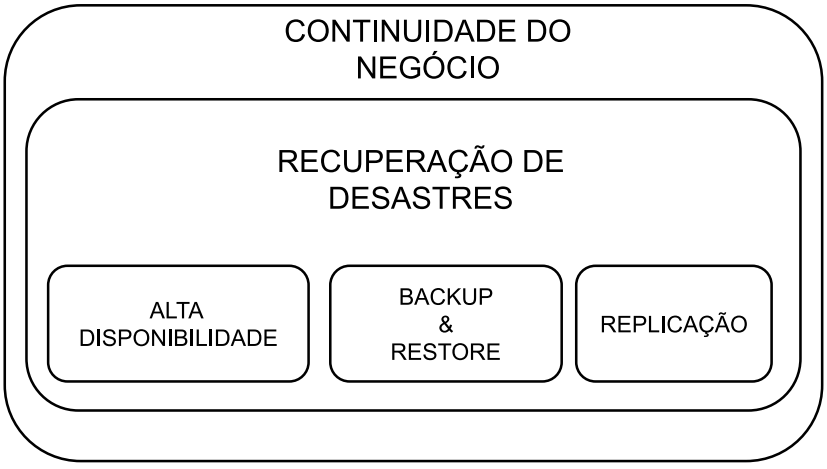


Figura 7-33 Continuidade dos Negócios

7.7.1. Conceitos Importantes

Recuperação operacional

A capacidade de reagir a interrupções nos serviços em nível de DATACENTER, normalmente causadas por corrupção de dados lógicos, erro humano ou falhas de hardware local.

Recuperação de desastres

A capacidade de reagir a uma interrupção local nos serviços e restaurar as funções essenciais aos negócios da organização, normalmente em um outro local.

Mobilidade de dados

A capacidade de criar réplicas para mover dados de um sistema para outro.

Os conceitos de *downtime*, alta disponibilidade e RPO e RTO são fundamentais para a compreensão das tecnologias e serviços relacionados com a recuperação a desastres.

Downtime

A **Figura 7-34** representa um incidente no tempo T1 ocorrido em um determinado sistema. O tempo decorrido entre T1 e T2 é o tempo de downtime. O tempo T0 é o tempo de perda de dados durante o downtime. Quanto menor o tempo de downtime e menor o tempo de perda de dados, maior a disponibilidade do sistema e maior o investimento em tecnologias que permitem a recuperação do site. A alta disponibilidade significa que o serviço de TI está continuamente disponível para o cliente havendo pouco downtime e recuperação rápida. Resumindo, downtime é o tempo que o sistema de TI fica fora do ar.

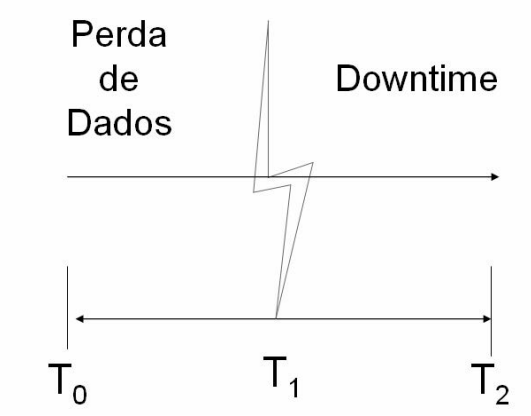


Figura 7-34 Downtime

Alta Disponibilidade

A alta disponibilidade pode acontecer no nível do hardware com a utilização de componentes redundantes e tolerantes a falhas. Os servidores podem ser tolerantes a falhas com o apoio da versão específica do sistema operacional. Alguns sistemas operacionais, como o Microsoft Windows Server 2008 DATACENTER e o Windows 2008 for Itanium-Based Systems, possuem o recurso de partição por hardware e permitem que os servidores sejam particionados dinamicamente. Este recurso permite que o Windows trabalhe em partições isoladas do hardware que podem ser modificadas com o upgrade de componentes com o sistema rodando.

A alta disponibilidade também acontece no nível do sistema operacional e no nível da aplicação. No nível do sistema operacional, os clusters de failover possibilitam a obtenção da alta disponibilidade. Um cluster de failover é um grupo de servidores independentes com dois ou mais discos compartilhados que trabalham juntos para aumentar a disponibilidade das aplicações e dos serviços. No nível da aplicação com o apoio do sistema operacional, pode-se ter a alta disponibilidade para a carga de trabalho específica.

A alta disponibilidade alcançada por um site ou mesmo um sistema é indicada por métricas. É muito comum associar o nível de alta disponibilidade à porcentagem de uptime do sistema. A fórmula abaixo permite saber o grau de disponibilidade de um sistema:

A=MTBF/(MTBF+MTTR)

Onde A é o grau de disponibilidade expresso em porcentagem, MTBF é o mean time between failures e o MTTR é o maximum time to repair. Se um sistema tem MTBF de 100.000 horas (mais que 11 anos) e o MTTR é de 1 hora, a disponibilidade (A) é de 99.9999 %. Ou seja, em 11 anos haverá 6 minutos de downtime.

Sistemas com cinco noves (99.999%) podem parar cinco minutos e 15 segundos ao longo do ano, considerando inclusive as paradas programadas. A **Tabela 7-4** ilustra a relação entre o nível de disponibilidade e o downtime.

Tabela 7-4 Disponibilidade e Downtime

DISPONIBILIDADE	DOWNTIME
99.9999%	32s
99.999%	5m e 15s
99.99%	52m e 36s
99.9 %	8h e 46m
99%	3 dias, 15 horas e 40 minutos

RPO e RTO

O RPO e o RTO são a tradução dos acordos de nível de serviço (SLA) para recuperação de desastres. A **Figura 7-35** ilustra duas medidas de tempo RPO e RTO importantes referentes à recuperação de sites e às operações de continuidade de negócios.

- **RPO (*recovery-point objective* – objetivo de ponto de recuperação)**: é o processo de olhar para trás que mede o quanto será necessário retroceder no tempo para obter uma boa cópia de reinicialização dos dados, se algo de ruim acontecer a eles. RPO é sobre perda de dados.
- **RTO (*recovery-time objective* – objetivo de tempo de recuperação)**: é o processo de olhar para frente que estima o tempo que levaria para seus negócios se tornarem totalmente operacionais novamente, uma vez que você tenha uma boa cópia de seus dados. Essa é uma medição para tempo de inatividade. RTO é sobre tempo de recuperação.

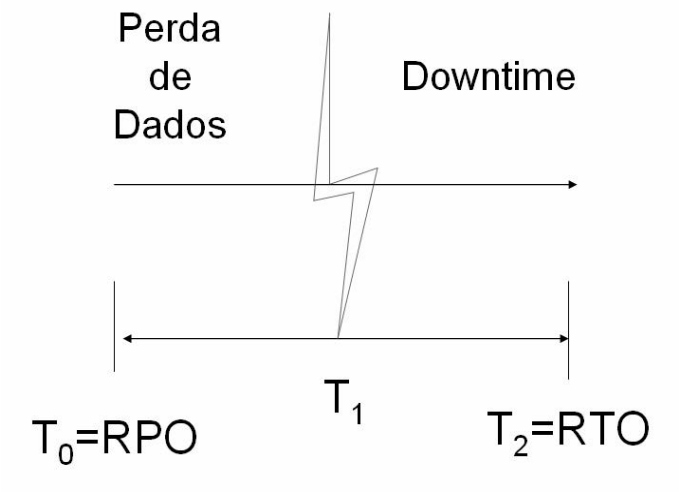


Figura 7-35 RPO E RTO

7.7.2. Clusters de Servidores

Os *cluster*s são grupos de servidores que trabalham juntos e que, para certas características, normalmente disponibilidade e desempenho, atuam como se fossem um único servidor.

O conceito de *cluster* é similar ao conceito de computação multiprocessada com a diferença de que a interconexão do *cluster* é realizada externamente, normalmente feita por uma rede privada, e a interconexão em sistemas de vários processadores é realizada internamente. Do ponto de vista do usuário da aplicação, quando um nó do *cluster* cai, o usuário só sente uma degradação temporária do serviço requisitado, mas não perde o serviço. A origem do *cluster* pode ser atribuída à invenção do *VAXcluster* para o sistema operacional VAX/VMS em meados da década de 80.

A ideia de utilizar *clusters* de servidores é aumentar a disponibilidade do ambiente, aumentar o desempenho da aplicação e reduzir a complexidade do gerenciamento. Diferentes modalidades de *clusters* endereçam diferentes objetivos.

Os softwares de VIRTUALIZAÇÃO trouxeram novas opções para a realização dos clusters. Pode-se agora obter alta disponibilidade de uma máquina virtual utilizando o próprio software de VIRTUALIZAÇÃO.

Principais Tipos de Clusters

Os *clusters* podem ser classificados em:

- **Clusters de alta disponibilidade (*high availability* – HA):** os clusters de alta disponibilidade endereçam redundância com capacidade de failover automático.
- **Clusters de balanceamento de carga (*load balancing* – LB):** os clusters do tipo load balancing endereçam a melhoria da capacidade para a execução da carga de trabalho (WORKLOAD).

- **Clusters de alta performance (HPC e HTC):** os clusters do tipo alta performance endereçam o aumento da perfomance da aplicação.
- **Grid:** os clusters em grid endereçam o melhor dos mundos com alta disponibilidade e alto desempenho. Podem ser globais, empresariais ou simplesmente grid clusters.

A escolha da tecnologia de *cluster* HA ou LB depende se os aplicativos a serem executados possuem estado de execução demorada na memória:

- O cluster HA destina-se aos aplicativos que têm estado de execução demorada na memória ou que têm estados de dados frequentemente atualizados. Esses são denominados aplicativos com monitoração de estado e incluem aplicativos de bancos de dados e aplicativos de mensagens. A utilização típica dos clusters de failover inclui servidores de arquivos, servidores de impressão, servidores de bancos de dados e servidores de mensagens.
- O cluster LB destina-se aos aplicativos que não têm estado de execução demorada na memória. Esses são denominados aplicativos sem monitoração de estado. Um aplicativo sem monitoração de estado trata cada solicitação do cliente como uma operação independente e, portanto, pode balancear a carga de cada solicitação de forma independente. Em geral, os aplicativos sem monitoração de estado possuem dados somente de leitura ou dados frequentemente alterados. Servidores web, VPNs, servidores FTP, firewalls e servidores proxy costumam usar o Balanceamento de Carga de Rede. Os clusters de Balanceamento de Carga de Rede também podem oferecer suporte a outros serviços e aplicativos baseados em TCP ou UDP.

A **Figura 7-36** ilustra a classificação dos *cluster*s baseada em requisitos de desempenho e disponibilidade.

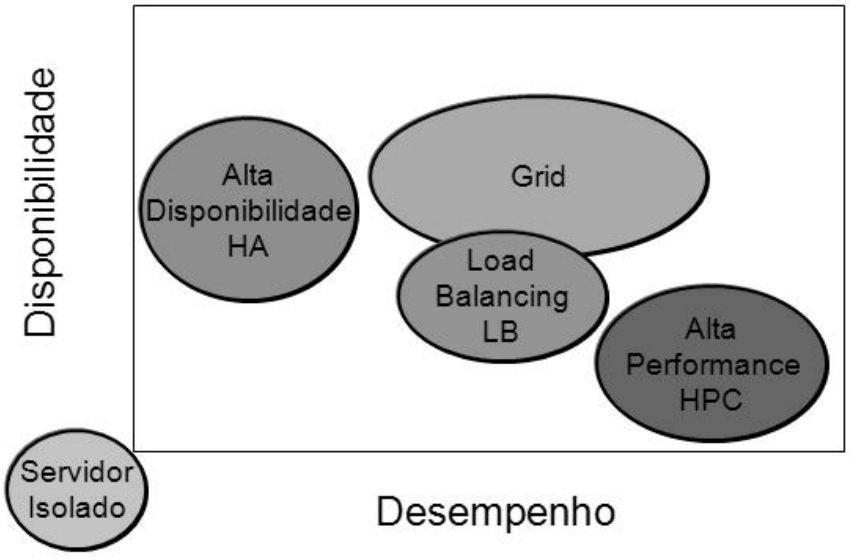


Figura 7-36 Classificação dos Clusters

Independentemente do tipo, todo *cluster* possui:

- **Múltiplos nós (físicos ou virtuais):** nos clusters de alta disponibilidade e nos clusters de alto desempenho um requisito especial é que os nós sejam homogêneos, possibilitando o gerenciamento como uma entidade única. Clusters do tipo load balancing e grid frequentemente possuem nós heterogêneos.
- **Interconexão:** a interconexão normalmente baseada em rede Ethernet/TCP-IP deve permitir a comunicação entre os nós do cluster.
- **Acesso aos dados e ao storage:** dependendo do tipo de cluster, os dados podem ser acessados por todos os nós do cluster (*share-everything cluster model*) ou os nós podem trabalhar isolados (*share-nothing cluster model*). Os clusters de alta disponibilidade e alto desempenho utilizam o modelo share-nothing. Um exemplo de share-everything cluster model é o cluster Oracle RAC.
- **Software de cluster:** o software de cluster providencia o protocolo para a comunicação entre os nós do cluster e normalmente é considerado um serviço do sistema operacional. Os dois principais softwares de cluster são o Microsoft Cluster Server (MSCS) e o Red Hat Cluster Suite.

Principais Clusters HA Disponíveis

- Microsoft Server Cluster (versões do Windows Server 2003 R2 e do Windows Server 2008 R2).
- Red Hat Cluster Suite (RHCS): oferece dois principais serviços de cluster com foco em alta disponibilidade.
- Symantec Veritas Cluster Server (VCS).
- Oracle Clusterware (versões do Oracle 10g, Oracle 11g).

7.7.3. Backup e Restore

O *backup* é uma cópia dos dados de produção, criada e retida para o propósito de recuperar dados deletados ou corrompidos. O *restore* é o processo de recuperação dos dados.

Os serviços de backup e restore em TI são parte de um processo mais amplo que visa garantir a disponibilidade do DATACENTER e da infraestrutura de TI e, portanto, garantir a continuidade do negócio.

O backup/restore necessita ser realizado por várias razões:

- Requisitos de negócio.
- Requisitos legais.
- Proteção contra falhas de hardware.
- Proteção contra falha da aplicação.
- Proteção contra erro do usuário.

- Recuperação de desastres.
- Atingimento de níveis de serviço específicos.

O aspecto chave para definir a forma de realizar o backup é a natureza dos dados objeto do backup. Deve-se definir também uma estratégia para recuperação dos dados do backup e testá-la antes da sua necessidade real. É comum acontecerem surpresas quando do uso do restore, ou seja, quando da tentativa de recuperar os dados. A própria forma de realizar o *restore* dos dados deve ser devidamente entendida e verificada com o fabricante do software de backup. A sugestão é que, quando do treinamento específico da ferramenta de backup a ser utilizada, inclua-se também o treinamento no *restore* do ambiente.

Normalmente pode ser feito um questionário com os donos dos dados para a realização de uma correta estratégia de backup. O questionário deve tentar descobrir quando deve ser feito o backup, quais dados devem ir para o backup, por quanto tempo devem ser mantidos e onde os backups devem ser feitos. Este procedimento deve fazer parte de um plano de continuidade do negócio mais amplo.

Nem todos os dados de uma organização precisam ir para backup ou necessitam de backup regularmente. Os sistemas operacionais de servidores, por exemplo, a menos que sofram algum tipo de patch (correção), não precisam de backup regular.

O processo de backup é ilustrado na **Figura 7-37**.

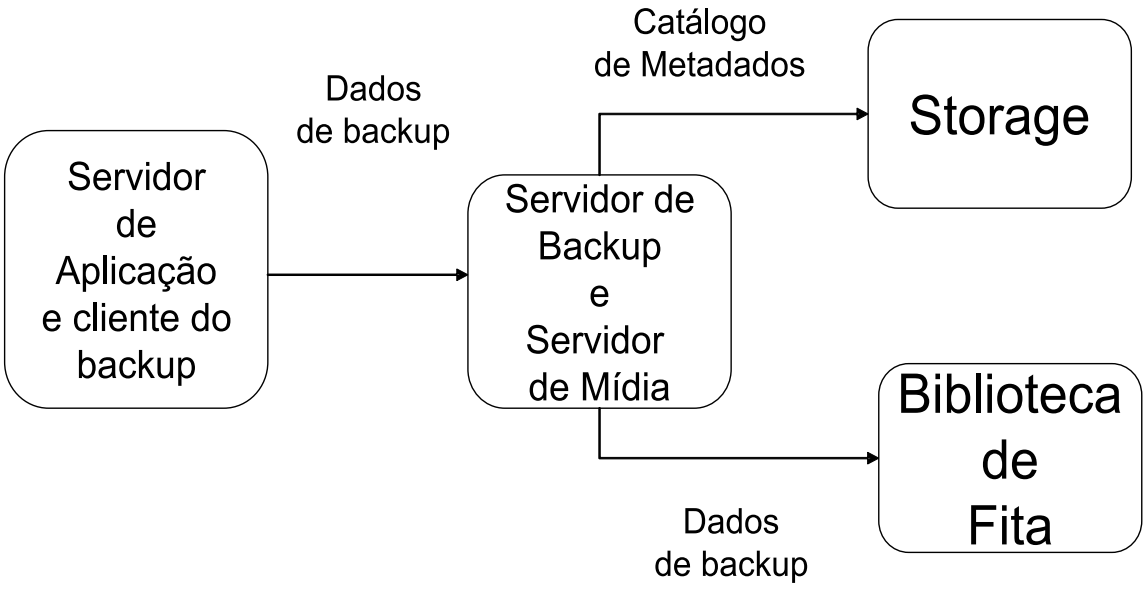


Figura 7-37 Processo de Backup

O aplicativo de backup, que normalmente é disponibilizado pelo fabricante, deve possibilitar realizar uma política de backup corporativa que envolva dados dispersos geograficamente e questões específicas dos sistemas operacionais e aplicativos que serão os focos do backup. Veja, estamos falando de aplicativos para o mundo empresarial. Esqueça o improviso quando se fala de backup e restore.

A unidade de fita é o dispositivo natural para a realização do backup devido ao seu custo benefício quando comparado a outras mídias. O padrão vigente de unidade de backup é o LTO-4, que, em unidades de quatro *drivers,* poderia fornecer uma taxa de 1.73 TB por hora. Um aspecto chave do backup baseado em fitas é a definição da política de rotação das fitas e o tempo de retenção delas. Um bom parâmetro de tempo de retenção on site é 30 dias e 180 dias para off-site, não considerando aqui imposições legais.

Outro aspecto relevante é a segurança do backup. Técnicas de encriptação podem ser utilizadas, mas degradam o desempenho do backup e do restore.

Backup em Disco e em Fita

A ideia de utilizar como parte da solução de backup também o backup em disco (disco-para-disco ou D2D) vem ganhado adeptos, permitindo combinar o que há de melhor em ambas as tecnologias (disco-para-disco-para-fita ou D2D2T). Discos e Fitas apresentam as seguintes características:

- **Disco**: velocidade, simplicidade.
- **Fita**: robustez, menor preço, opção de recuperação de desastre.

Existem muitas dúvidas sobre se realmente compensa ter uma estratégia de backup/restore que envolva discos e fitas. Um estudo realizado pelo Clipper Group comparou o custo do backup realizado em fita LTO-3 com o custo do backup em disco SATA. O estudo concluiu que o backup em disco tem um custo de aquisição 6.5 vezes maior do que o do backup em fita. Além disso, o espaço físico ocupado pelo backup utilizando discos no *storage* é 6.2 vezes maior do que o espaço ocupado pela unidade de fita, e o custo com energia e refrigeração é 25 vezes maior quando se utiliza backup em disco. Pode-se concluir que o backup em disco deve fazer parte de uma estratégia mais ampla de backup onde se busca velocidade, em um primeiro momento, através do backup em disco, e robustez, em um segundo momento, com a unidade de fita.

A utilização de discos de baixo custo em relação à unidade de fita para backup oferece as seguintes vantagens, segundo a CommVault:

- Recuperação de dados mais rápida.
- Os discos são, normalmente, mais rápidos do que as fitas, especialmente quando se leva em consideração o tempo de montagem e de busca dos dados em fitas.
- Os problemas com falha nas fitas são eliminados.
- Os discos têm acesso aleatório e são otimizados para pesquisas – as fitas têm acesso sequencial e podem ser muito mais lentas para buscas aleatórias de arquivos.
- Múltiplos servidores podem acessar os dados simultaneamente – os discos têm várias cabeças de leitura/gravação.
- O uso de discos elimina os erros humanos típicos no manuseio das fitas.

Uma opção natural hoje é a de utilizar uma combinação de discos e fita para backup. Esta opção é diferente da opção VTL (*virtual tape library*) para o backup em disco. O

VTL emula uma unidade de fita em uma unidade de discos. O VTL se mostrou de pouca escala pelo fato de ter uma interface isolada. Também é necessário utilizar um software adequado de gerenciamento de backup. Os discos utilizados pela VTL para o *storage* de backup são normalmente do tipo SATA para reduzir o custo da solução.

Em redes SAN é natural que a unidade de backup seja inserida na própria rede SAN. Para isto é necessário verificar se a unidade de backup, normalmente uma biblioteca de fitas, possui a interface correta para ser instalada na rede SAN. Também o licenciamento do software para este caso precisa ser cuidadosamente planejado, pois os servidores da SAN precisam estar devidamente licenciados para a utilização do software de backup, além do licenciamento dos agentes necessários. Cada fabricante de software de backup tem uma forma exclusiva de licenciamento e que muda constantemente.

A **Figura 7-38** ilustra uma típica arquitetura de backup.

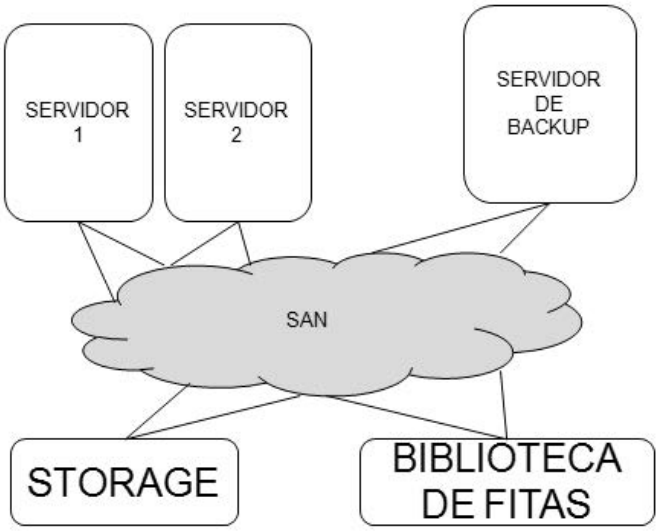


Figura 7-38 Arquitetura de BACKUP em SAN

Quando se planeja uma estratégia de backup e recuperação para um ambiente de servidor virtualizado, há diversos fatores a serem considerados. Deve-se considerar os vários tipos de backups que podem ser feitos, o estado da máquina virtual. Também é necessária uma estratégia para o backup das máquinas virtuais que passa pela integração de uma ferramenta de backup adequada com o aplicativo de backup disponibilizado pelo software de VIRTUALIZAÇÃO normalmente baseado em snapshot das máquinas virtuais.

7.7.4. Desduplicação

A desduplicação de dados é uma nova tecnologia para gerenciar o crescimento exponencial de dados e fornecer proteção. Para eliminar dados redundantes do *storage*, essa técnica de backup salva uma única cópia de dados idênticos e substitui todas as outras por indicadores que apontam para essa cópia. Desduplicar não é comprimir os dados.

A desduplicação examina a redundância dos dados que vão para backup. Em muitos casos é possível sair de um backup de 15 TB para 1 TB utilizando a desduplicação. A taxa de redução do backup vai depender do tipo de dado a ir para backup, dos períodos de retenção, da frequência dos backups completos e da tecnologia de desduplicação utilizada. É possível atingir reduções de até 500:1 (99,8%).

A desduplicação permite:

- Economias com menor investimento em discos.
- Melhor utilização da capacidade.
- Menos dependência de backup em fita.
- Recuperação mais rápida após uma pane.

Os aplicativos de backup são um componente essencial das estratégias de backup e recuperação de desastres. Da forma que era pensado o backup anteriormente, quase sempre as mídias utilizadas armazenam cinco a dez vezes mais dados do que pode ser encontrado no armazenamento primário do mesmo DATACENTER. Como os dados continuam crescendo sem parar e mais atenção é dirigida à proteção eficiente dos dados e à recuperação de desastres, as abordagens tradicionais ao backup não são mais aceitáveis. As soluções de backup e desduplicação baseadas em disco estão ficando obrigatórias em várias abordagens de gerenciamento de dados e agora se somam aos backups baseados em unidades de fita. Os novos sistemas de desduplicação se integram facilmente ao software de backup existente, permitindo obter maior eficiência na execução e recuperação do backup.

7.7.5. Arquivamento (Archive)

As organizações atuais têm necessidades legais e regulamentares referentes a arquivos de dados que precisam estar bem organizados e armazenados. Registros específicos devem ser localizados e recuperados rapidamente e de maneira eficiente. Acontece que as aplicações de backup são concebidas para a recuperação dos dados de eventos não planejados, como uma falha acidental de um dispositivo ou mesmo a exclusão de um determinado arquivo.

Os backups são cópias secundárias da informação primária. Eles fornecem proteção de curto prazo para os dados de produção para garantir a continuidade das atividades. Eles estão disponíveis e são gerados e substituídos sistematicamente. A principal característica do backup é a capacidade de fornecer uma cópia da informação no tempo adequado, a fim de proteger os processos empresariais críticos. Por este motivo, o backup é a solução adequada para a continuidade empresarial e recuperação de desastres.

O arquivamento (*archive),* por outro lado, trata de cópia principal das informações. As informações arquivadas são valiosas e conservadas para referências futuras e devem ter garantias de autenticidade. O archive trata de informações normalmente em sua forma final e sujeitas a pouca ou nenhuma modificação. Como passam a ser a única cópia das informações, devem ser mantidas por períodos de longo prazo (meses, anos ou décadas). O archive é focado no acesso e na recuperação de um trecho específico de

informação, em vez de todo o conteúdo (como o backup). Processos de archive costumam possuir informações específicas como prazos, incluindo a sua supressão.

Novas demandas, principalmente as regulamentadas (*compliance),* requerem a disponibilização de dados históricos online. Se o backup utilizar os meios tradicionais para estes arquivos vivos, o que acontece é que os dados são enviados para backup na unidade de fita e somem do *storage*. Qualquer demanda de consulta destes dados requer uma grande intervenção dos administradores de backup, pois estes precisam ser recuperados da fita e os tempos envolvidos nesta operação são normalmente proibitivos.

A **Figura 7-39** ilustra a diferença entre uma operação tradicional baseada em backup e outra baseada em *archive*. Observe que nem todos os dados vão para backup na opção com *archive*. Dados que precisam estar vivos e sofrem pouca modificação podem estar em archive online para serem consultados.

Na **abordagem tradicional** o crescimento de dados do ambiente de produção deve ser absorvido por um dispendioso armazenamento principal baseado em *storage*. Conforme crescem os requisitos do *storage* principal, crescem os requisitos de backup e arquivamento.

Na **abordagem baseada em** *archive* o arquivamento passa a ser uma extensão do ambiente de produção, permitindo o arquivamento mais cedo e a redução das necessidades de *storage* principal e de backup. As restaurações são mais rápidas e o SLA dos dados em *archive* em geral é melhorado.

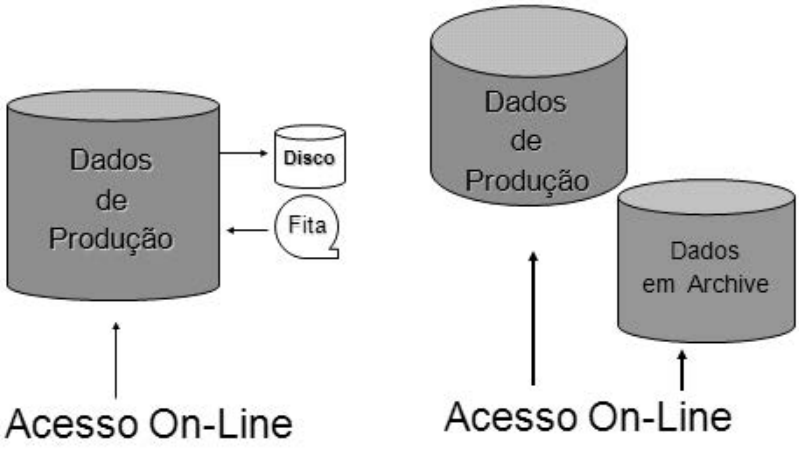


Figura 7-39 Archive

Uma boa solução de backup-restore para DATACENTER deve contemplar aspectos como, por exemplo, gerenciamento centralizado de jobs, janela de backup reduzida, agentes para múltiplos sistemas operacionais, ambientes virtualizados, contemplar a desduplicação de arquivos e dados, aceitar diferentes mídias de armazenamento e ofertar proteção contínua local e remota. A eficiência desses serviços depende da correta integração e do conhecimento e alinhamento dos processos de backup, recuperação e arquivamento.

7.7.6. Replicação

A replicação é a tecnologia normalmente utilizada para proteção de dados e recuperação de desastres, mas também pode ser utilizada para replicação de dados dentro do mesmo ambiente físico.

Importante ressaltar que:

- Replicação não é mobilidade dos dados. A mobilidade de dados trata da capacidade de mover dados de um sistema para outro. A replicação é a cópia de dados de um sistema e seus discos para outro sistema com um conjunto independente de discos.
- Replicação também não é o mesmo que espelhamento, pois o espelhamento trata ambos os conjuntos de discos como um único volume lógico. O espelhamento está confinado a um único sistema e a replicação move os dados de um sistema para outro.

O resultado final da replicação é a obtenção de dois conjuntos de dados consistentes e iguais em dois lugares diferentes. Importante salientar que deverá haver espaço suficiente para absorver o volume replicado e que, muitas vezes, o nível de RAID utilizado no storage nos dois conjuntos de discos é diferente.

Existem diversas formas de fazer a replicação, indo desde a replicação pelo servidor (*host based*), replicação pelo storage (*array based*) e até a replicação pela rede (*appliance based*), que permitem que as empresas desenvolvam uma abordagem coesa em relação à proteção de dados concentrada principalmente na recuperação dos dados, não apenas na realização de tarefas de backup.

A replicação, em resumo, pode ocorrer basicamente de três maneiras:

- Baseada em software de servidor (*Host Based*).
- Baseada em appliance de rede (*Appliance Based*).
- Baseada em array (*Array Based*).

A replicação pode limitar a exposição a tempos de inatividade planejados e não planejados, possibilitando operações contínuas. Normalmente uma organização precisa de uma replicação eficiente de dados para cumprir os padrões corporativos ou governamentais e, ao mesmo tempo, continuar atendendo aos requisitos de baixo custo total de propriedade. Além disso, a solução de replicação deve ser flexível para se ajustar à mudança imposta pelo ambiente.

Replicação Síncrona e Assíncrona

A replicação remota pode ser classificada em dois tipos:

Replicação Síncrona: de acordo com a **Figura 7-40** (a):

1. O servidor escreve o dado na fonte.
2. O dado da fonte é replicado para o destino no site remoto.

3. O *ack* do destino é enviado para a fonte.
4. O *ack* da fonte é enviado ao servidor.

A replicação síncrona propicia excelente integridade aos dados, mas exige alto desempenho da rede. Suas principais características são:

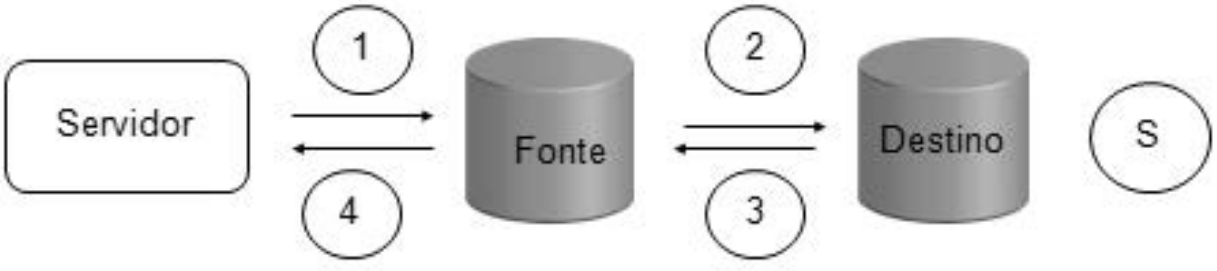
- RPO de zero ou próximo de zero segundos.
- As duas imagens são idênticas.
- Exige grande largura de banda de rede.
- Em geral é realizada de um site principal para um ou dois secundários em distâncias menores do que 200KM.

Replicação Assíncrona: de acordo com a **Figura 7-40** (b):

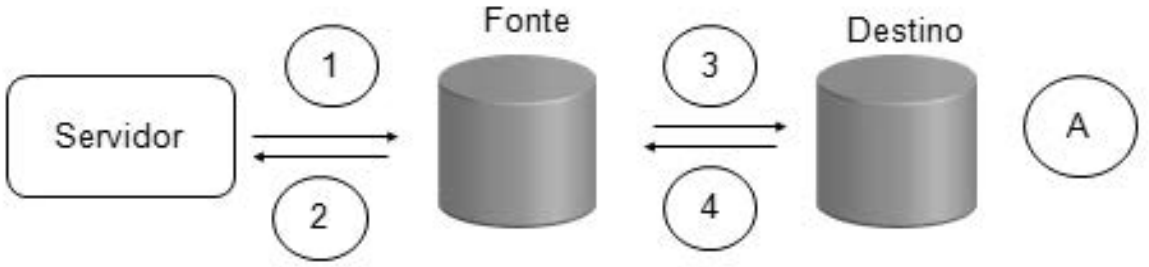
1. O servidor escreve o dado na fonte.
2. O *ack* é enviado imediatamente ao servidor.
3. O dado é transmitido ao destino no site remoto mais tarde.
4. O *ack* do destino é enviado de volta à fonte.

A replicação assíncrona propicia menor integridade dos dados, mas possibilita o uso de redes não performáticas na conexão entre os sites. Suas principais características são:

- RPO de 30 minutos a horas.
- Destino atualizado periodicamente.
- Distância ilimitada.
- Cópia reinicializada no site secundário se a sessão falhar.
- Otimizada para pequena largura de banda.
- Em geral um site principal para um site secundário.



(a)



(b)

Figura 7-40 Replicação Síncrona (a) e Assíncrona (b)

Replicação baseada em snapshots/clone ou CDP/CCR

A forma de realizar a replicação evoluiu e pode ser basicamente dividida em dois grandes grupos:

- **Replicação baseada em snapshots ou clones:** replicação similar a uma máquina fotográfica.
- **Replicação baseada em CDP (*continuous data protection*) ou CCR (*continuous remote replication*):** replicação similar a uma câmera de vídeo – os dados podem ser facilmente reconstruídos de qualquer ponto. CDP utiliza uma replicação local síncrona e CCR seria a replicação remota contínua, normalmente assíncrona.

Dentro destes grupos existem vários métodos de replicação disponíveis, cada um com um perfil diferente e a chave para determinar qual solução de replicação é ideal para determinado aplicativo é compreender os níveis de serviço realmente necessários.

Ao investigar diferentes soluções de replicação, percebe-se que cada método – seja ele síncrono ou assíncrono – exige a discussão e a compreensão de vários aspectos essenciais. São eles:

- O impacto no tempo de resposta do aplicativo de produção.
- A infraestrutura: que equipamento adicional será necessário para dar suporte ao processo de replicação?
- Os links de comunicação: qual o tamanho e o custo do link de comunicação necessário para dar suporte ao processo?
- Quais os RPO e RTO aceitáveis?

A VIRTUALIZAÇÃO simplifica e reduz o custo dos processos de alta disponibilidade e recuperação de desastres.

A VIRTUALIZAÇÃO simplifica a replicação, na medida em que permite a integração do software de VIRTUALIZAÇÃO com o storage, automatizando o processo. Além disso, permite grande economia na construção do site de recuperação a desastres, pois otimiza o uso de servidores e unidades de armazenamento.

7.7.7. Conceito na Prática: EMC RecoverPoint

O EMC RecoverPoint é uma solução única que oferece as vantagens das soluções baseadas em hosts e em arrays, enquanto replica seus dados de qualquer array baseado em SAN para qualquer outro array baseado em SAN em sua rede Fibre Channel ou IP atual – de modo rápido, simples e econômico – usando uma combinação de opções de divisão de gravação com base em host, no EMC CLARiiON ou em malha inteligente.

O EMC RecoverPoint/SE é uma solução para a série de arrays EMC CLARiiON, e o EMC RecoverPoint é uma solução abrangente para o DATACENTER. Os dois produtos fornecem replicação local síncrona usando CDP (Continuous Data Protection, proteção contínua dos dados) e CRR (Continuous Remote Replication, replicação remota contínua) assíncrona com recursos de recuperação em qualquer point-in-time. A família RecoverPoint protege empresas contra a perda de dados resultante de problemas comuns, como falhas no servidor, corrupção de dados, erros de software, vírus e erros de usuário final, ao mesmo tempo em que protege contra eventos catastróficos que possam provocar a paralisação de todo o DATACENTER.

O RecoverPoint contém um módulo para realizar o CDP local e um módulo para realizar o CRR remoto. Esses módulos podem ser executados no mesmo utilitário do RecoverPoint para fornecer proteção de dados CLR (*Concurrent Local and Remote*, simultânea local e remota), o que proporciona uma solução única que protege os mesmos dados tanto local quanto remotamente. Isso permite que você reduza custos e facilite o gerenciamento com base em suas necessidades exclusivas de proteção completa dos dados.

A CDP do RecoverPoint rastreia alterações nos dados no nível de bloco e as registra em log usando a tecnologia CDP. Esse recurso é utilizado para proteger dados críticos contra falhas físicas ou lógicas, como paralisações de servidor, corrupção de dados, erros de software, vírus e erros comuns de usuários. Isso complementa as soluções atuais de backup e ajuda a eliminar a necessidade de manter rígidas janelas diárias de backup. As tecnologias avançadas de compactação de dados geram uma redução inigualável do armazenamento necessário para registro, permitindo que mais dados sejam mantidos online para recuperação mais rápida.

A CRR do RecoverPoint é uma solução corporativa de replicação de dados que oferece recuperação point-in-time, permitindo a proteção total contra desastres locais. Desempenho de nível corporativo, capacidade de expansão e recuperação instantânea são combinados para garantir a consistência dos dados com uma recuperação em segundos ou minutos e não em horas ou dias. A CRR do RecoverPoint fornece replicação síncrona e replicação bidirecional contínua e assíncrona, sem limitação de distância e no nível de bloco. Sua tecnologia foi projetada para manter a consistência de dados dependente da ordem de gravação e alterna dinamicamente entre os modos de replicação síncrona e assíncrona com base nas políticas do usuário para desempenho e latência. Seus recursos avançados de redução da largura de banda e compactação de dados foram projetados para reduzir drasticamente os requisitos de largura de banda da WAN e os custos associados.

O RecoverPoint usa tecnologia patenteada e comprovada para reduzir drasticamente os requisitos de largura de banda para replicação, por meio de compactação de dados e tecnologias inteligentes de redução de largura de banda que proporcionam uma diminuição de até 10:1 na largura de banda da WAN. Essas tecnologias inteligentes permitem que os clientes contem com o melhor nível possível de proteção para a largura de banda disponível, reduzindo drasticamente os custos da WAN, sobretudo em longas distâncias.

O RecoverPoint preserva a ordem de gravação dos dados alterados e armazena-os com bookmarks específicos ao horário e, opcionalmente, ao aplicativo, permitindo a recuperação imediata em qualquer point-in-time, sem perda de dados. Ele pode recuperar aplicativos rapidamente em um período ou evento selecionado com consistência garantida, mesmo que os dados abranjam diversos volumes e servidores de armazenamento heterogêneo.

O RecoverPoint permite que os dados dos aplicativos sejam acessados em qualquer point-in-time. Para isso, seleciona marcadores inteligentes específicos ao horário ou ao aplicativo. Os dados de aplicativos no point-in-time selecionado são recuperados instantaneamente e podem ser lidos e gravados imediatamente pelo host. O RecoverPoint permite acesso de leitura/gravação a dados replicados sem afetar o processo de replicação. Na recuperação, isso permite que os dados sejam testados em diversos points-in-time para determinar o melhor ponto de início da recuperação. Esse recurso pode ser utilizado para desafogar backups, permitir testes de aplicativos em tempo real, dar suporte à recuperação sob demanda, migrar dados e muitas outras finalidades valiosas de processamento de dados. O RecoverPoint também aceita vários pontos de recuperação consistentes com transações – todos indexados em um point-in-time específico.

O RecoverPoint garante uma réplica consistente de dados de negócios, caso ocorra uma possível falha ou desastre. O RecoverPoint sempre mantém a consistência, mesmo que os dados abranjam diversos sistemas e servidores de armazenamento heterogêneo.

O RecoverPoint oferece integração com aplicativos de negócios da Microsoft, SAP, Oracle e outras, para facilitar a proteção e a recuperação inteligentes, reduzindo drasticamente o tempo de recuperação de aplicativos e eliminando a perda de dados. A família RecoverPoint monitora os eventos de aplicativos e de ambiente durante o processo de replicação. Os eventos importantes que ocorrem no aplicativo, no host, na rede ou no nível de armazenamento são registrados e é feita uma análise de causas básicas para examinar quaisquer falhas resultantes do processo.

7.7.8. Replicação com Desduplicação

A desduplicação também permite otimizar a replicação com a redução da largura de banda WAN necessária para a transferência dos dados. A desduplicação também permite reduzir o tempo necessário para criar várias cópias de backups para fins de consolidação em fita ou recuperação a desastres. A ideia central é remover os dados redundantes da replicação antes de realizá-la, como é feito para o backup.

7.8. Desktop e VIRTUALIZAÇÃO

A gerência dos *desktops* em qualquer organização sempre foi um desafio. O uso do conceito de TCO em TI foi validado pelo cálculo do custo de propriedade do desktop baseado em Windows. Na época, era comum os grandes fabricantes utilizarem-se desta nova maneira de pensar o custo para justificar a aquisição por parte de clientes de desktops de grife, que tinham melhor gerenciamento e MTBF e, portanto, deveriam ser mais econômicos ao longo do ciclo de vida do equipamento.

A computação distribuída espalhou a execução da aplicação ou parte dela em desktops isolados ou conectados na rede para atender à demanda da organização, mas se mostrou cara e muitas vezes de difícil gerenciamento. A alteração da geografia de muitas organizações para maior complicou mais ainda o upgrade das aplicações e outras tarefas que precisavam ser executadas no desktop localmente, aumentando o custo do suporte.

Surgiu então a ideia do Windows Server Based Computing (WSBC), concebida pela Microsoft e pela Citrix de consolidar o gerenciamento e a execução das aplicações de desktop no servidor, mas mantendo a flexibilidade do ambiente Windows do desktop. A Citrix, com o seu Presentation Server, e a Microsoft, com o Windows Terminal Services, foram líderes nesta solução baseda na agregação das sessões de vários usuários dentro de um único SO e fizeram emergir o conceito de *thin client* no mundo x86.

O *thin client* funciona sem SO local ou aplicativos para instalar e reparar, assim os custos de manutenção e suporte são reduzidos, possibilitando aos usuários uma experiência semelhante à do desktop. Os aplicativos são executados no servidor, o qual leva as telas em tempo real aos usuários que utilizam desktops *thin client*. Segundo a Wyse, empresa líder no fornecimento de *thin client*, esta solução funciona muito bem para 80% dos aplicativos.

Acontece que a funcionalidade multiusuário do Windows em certas situações se mostrou inadequada pelo simples fato de o Windows não ter sido projetado para funcionar assim. Por exemplo, no Windows Server Based Computing, se aplicações intensivas em CPU forem usadas simultaneamente por poucos usuários, os outros usuários pendurados no servidor terão o desempenho degradado também para a execução de suas aplicações.

Depois surgiu o *thin computing streaming*, uma tecnologia que o sistema operacional e os aplicativos são entregues pouco a pouco, on demand, a um *desktop* stateless (sem partes móveis) altamente seguro, onde são executados localmente na CPU de cada usuário. Os aplicativos rodam localmente, como ocorre no atual modelo de desktop FAT, e poupam banda.

Recentemente surgiu o VDI (*Virtual Desktop Infrastructure*) baseado em VIRTUALIZAÇÃO. O VDI permite que as aplicações dos usuários rodem em máquinas virtuais isoladas e ao mesmo tempo compartilhem recursos de hardware como CPU, memória, disco e rede do servidor. A WYSE denomina de Enterprise Desktop Virtualization (EDV) a combinação da virtualização com o *thin client*.

No VDI cada usuário roda sua aplicação em seu próprio SO reduzindo as chances de que outro usuário possa interferir na sua execução. O VDI separa o hardware do software e o HYPERVISOR encapsula a aplicação e o SO em uma máquina virtual que roda em um servidor do DATACENTER.

Normalmente o protocolo de conexão remoto entre o desktop do usuário e o servidor é o *Remote Desktop Protocol* (RDP) da Microsoft. Outros protocolos de desktop utilizados são o *Independent Computing Architecture* (ICA) e o *Virtual Network Computing* (VNC).

7.9. Gerenciamento e VIRTUALIZAÇÃO

A VIRTUALIZAÇÃO introduz um elemento que se destina a cuidar do gerenciamento das máquinas virtuais. O gerenciamento normalmente é realizado por um componente da solução de VIRTUALIZAÇÃO que requisita uma infraestrutura específica para o gerenciamento.

A ideia do gerenciamento unificado para todos os componentes do DATACENTER é ainda um sonho. Os fabricantes utilizam padrões distintos e cada componente possui seu próprio console de gerenciamento. A proposta dos frameworks de gerenciamento é fornecer uma única plataforma de gerenciamento para todos os componentes, ainda uma ideia que precisa virar realidade.

7.10. Segurança e VIRTUALIZAÇÃO

A segurança da VIRTUALIZAÇÃO é um aspecto essencial a ser considerado quando da escolha do software de VIRTUALIZAÇÃO. Normalmente a segurança é garantida pelo HYPERVISOR.

O HYPERVISOR possui mecanismos de detecção de intrusão e controla o acesso aos discos e à rede.

A arquitetura do HYPERVISOR possibilita normalmente:

- Ter uma camada de VIRTUALIZAÇÃO que acaba contribuindo para a segurança.
- Compatibilidade com as técnicas de SAN, incluindo LUN masking e zoning.
- Implementação de técnicas de segurança de rede como:
  - Integração com AD.
  - Permissões e regras customizadas.
  - Controle de acesso do pool de recursos e delegação.
  - Controle das sessões de usuário do virtual center.
  - Trilhas de auditoria.

Diversos mecanismos de segurança estão sendo incorporados aos softwares de VIRTUALIZAÇÃO, melhorando a segurança do ambiente como um todo.

7.11. Conceito na Prática: Cisco DATACENTER 3.0

O Cisco DATACENTER 3.0 representa a estratégia da Cisco para facilitar a transformação do DATACENTER em um modelo virtualizado e integrado. O Cisco DATACENTER 3.0 combina uma arquitetura inovadora com um *roadmap* de tecnologia integrado. O projeto utiliza melhores práticas e serviços para ajudar os clientes a transformarem a infraestrutura, os processos e a organização do DATACENTER na busca de melhoria da eficiência e da minimização dos riscos, dos custos ou de possibilidades de ruptura.

7.11.1. Portfólio

O portfólio de tecnologia para o Cisco DATACENTER 3.0 envolve:

- **Infraestrutura de Rede**: propicia a conectividade física e lógica e o acesso a recursos de processamento e armazenamento físicos e virtualizados.
- **Infraestrutura de Serviços**: serviços que residem na plataforma de infraestrutura para assegurar a escalabilidade, gerenciabilidade e confiabilidade das aplicações e informações armazenadas no DATACENTER.
- **Serviços de Aplicação**: serviços focados em melhorar o desempenho e a disponibilidade das aplicações. Incluem serviços de aplicação e serviços de aplicação de longo alcance.
- **Ferramentas de automação e provisionamento**: visam reduzir a complexidade do DATACENTER.

7.11.2. Roadmap

O Roadmap descrito sinaliza como a Cisco enxerga a evolução do DATACENTER, saindo da consolidação para estruturas públicas chamadas de mercado.

1. **Consolidação**: o desenvolvimento de uma arquitetura de rede para o DATACENTER permite a melhoria da eficiência operacional, melhor utilização e disponibilidade da aplicação.
2. **Virtualização**: a introdução do *unified fabric* otimiza a orquestração e a hospedagem dos recursos de servidores, storage e rede virtualizados, melhorando a disponibilidade da aplicação e o tempo de resposta da TI.
3. **Automação**: o desenvolvimento da arquitetura de computação unificada permite a simplificação da infraestrutura através de uma arquitetura de VIRTUALIZAÇÃO pré-projetada, ativos de TI com ciclo de vida maior e escalabilidade adequada aos requisitos de mudança.
4. **Utility**: uma nuvem empresarial interna baseada nos modelos de arquitetura anteriormente construídos possibilita que os DATACETENTERS ofereçam classes de serviço em um modelo pague-pelo-uso.
5. **Mercado**: extensão das nuvens internas para operar inter-clouds através de DATACENTERS e provedores de serviço terceirizados. Pode-se mover serviços de dados para o ambiente de hospedagem mais apropriado em termos de requisitos de negócios e forças de mercado.

7.11.3. Componentes

Cisco Unified Computing System (UCS)

O Cisco UCS é a plataforma do DATACENTER 3.0 da Cisco. O sistema UCS integra uma rede unificada em 10 Gigabit Ethernet com uma arquitetura de servidores XEON x86, em um único domínio gerenciável. Os servidores blade da Cisco instalados no chassi UCS 5108 têm acesso ao *fabric* através de adaptadores que propiciam até 40 Gbps de banda por servidor blade. O *unified fabric* simplifica o cabeamento eliminando a necessidade de múltiplos adaptadores de rede e de storage com cabeamentos separados. O *unified fabric* suporta Ethernet e também FCoE, que podem ser gerenciados de forma independente com gerenciamento de banda e interferência entre classes de tráfego.

Os componentes da solução UCS são:

- **Cisco UCS 6100 Series Fabric Interconnect:** switches Ethernet de 10Gbps que consolidam o I/O dentro do sistema. Atualmente são duas opções de 20 e 40 portas que propiciam conectividade para Fibre Channel e Ethernet.
- **Cisco UCS 5100 Series Blade Server Chassis:** suportam até 8 blade servers e até dois fabric extenders em 6RU.
- **Cisco UCS 2100 Series Fabric Extenders:** propiciam até 4 conexões de 10Gbps entre cada blade server e o fabric de interconexão.
- **Cisco UCS B-Series Blade Servers:** blades baseados nos processadores Intel XEON 5500 e 5600. A tecnologia da Cisco para expansão de memória aumenta a performance e a capacidade para grandes cargas de trabalho.
- **Cisco UCS Manager:** possibilita o gerenciamento centralizado criando um domínio de gerenciamento unificado.
- Existem também servidores em RACK C-Series e adaptadores de rede que completam o portfólio.

A **Figura 7-41** ilustra o sistema Cisco UCS.

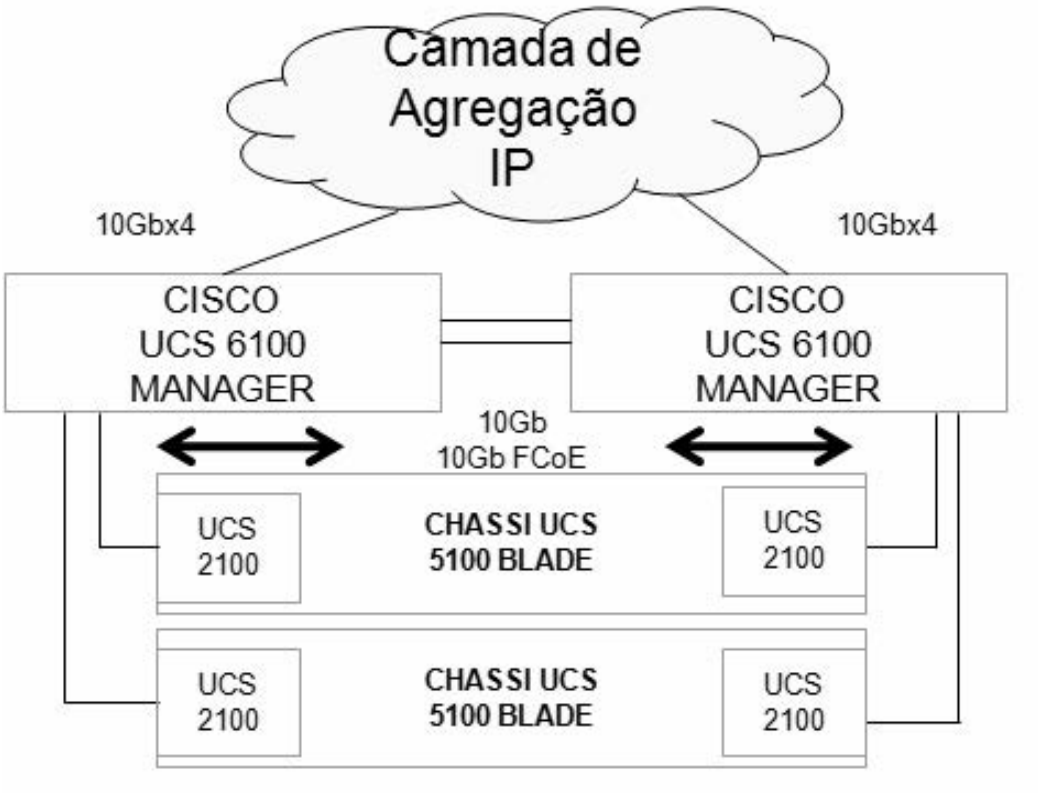


Figura 7-41 Cisco UCS

Cisco VN-LINK

Cisco VN-Link suporta links virtuais de rede entre cada máquina virtual e o *fabric* de interconexão, facilitando o gerenciamento e o movimento de máquinas virtuais em suas redes, enquanto mantém as características de rede, incluindo segurança. A camada de acesso tem sido fragmentada em três camadas em soluções que utilizam blades, o que traz implicações ruins como o aumento da latência nas relações VM-para-VM.

- Switches de acesso.
- Switches que residem nos servidores blade.
- Switches de software.

A **Figura 7-42** ilustra as três opções sugeridas.

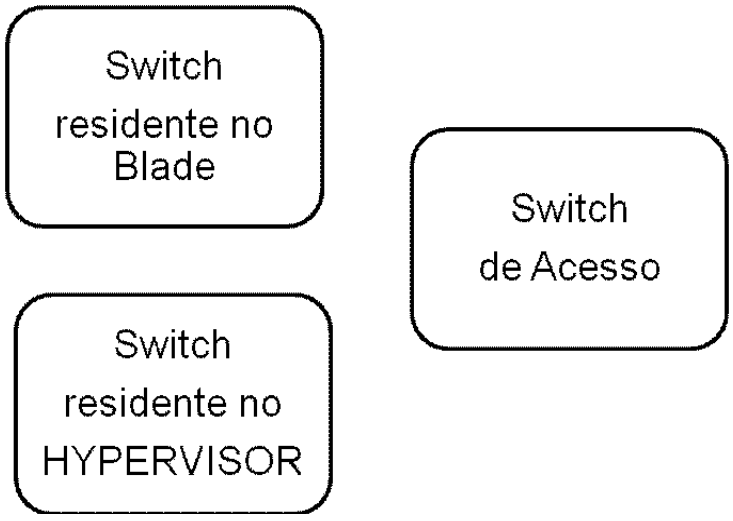


Figura 7-42 Cisco UCS e camada de acesso

As opções trazem flexibilidade, mas criam obstáculos para o efetivo gerenciamento. Cisco UCS simplifica reduzindo os switches envolvidos na comunicação VM-para-VM em uma interconexão fabric única.

Cisco UCS Manager

Transforma os recursos do Cisco UCS em um único sistema que simplifica a forma de estabelecer pools de recursos para ambientes virtualizados. O Cisco UCS possui uma interface GUI, interface em linha de comando e uma API que possibilita o UCS manager ser utilizado com softwares de terceiros.

Templates para perfis de serviço facilitam o processo de incorporar novos servidores no sistema, pois podem ser incorporados como pools de recursos em minutos.

Cisco UCS Tecnologia Estendida de Memória

A Cisco UCS tecnologia estendida de memória é uma opção para aumentar a densidade de máquinas virtuais sem aumentar o TCO. Esta tecnologia é disponibilizada para servidores blade do tipo Cisco UCS B250 M1 Extended Memory Blade Server e Cisco UCS C250 M1 Extended Memory Rack-Mount Server. A tecnologia permite mapear quatro memórias físicas em uma memória lógica. Este processo permite estender a memória dos servidores, melhorando a performance de soluções virtualizadas. Até 48 slots DIMM em uma solução estendida pode dar a organização excepcional flexibilidade no processamento, na capacidade de memória e no custo.

7.12. Referências Bibliográficas

10 Gigabit Ethernet Unifying Fabric: Foundation for the Scalable Enterprise, Dell Power Solutions, server Virtualization: Cisco Unified Computing System, Cisco Systems, 2009.

Blades Platforms and Network Convergence, Blade.org, 2008.

Cisco DATACENTER 3.0 At a Glance. Cisco Systems, 2009.

Cisco UCS Managem. At a Glance, Cysco Systems, 2010.

Cisco Unified Computing System. At a Glance. Cisco Systems, 2010.

Cisco VN-Link Virtualization-Aware Networking, Cysco Systems, 2009.

Clark, Tom. Designing Storage Area Network: A practical reference for Implementing Fibre Channel and IP SANs, Addison-Wesley, Second Edition, 2003.

Clark, Tom. IP SANs: A guide to iSCSI, iFCP and FCIP Protocols for Storage Area Networks, Addison Wesley, 2004.

DATACENTER Blade Server Integration Guide, Cisco, 2006.

Dell Business Ready Configuration for DATACENTER virtualization, Dell, 2009.

DELL Equallogic PS Series Architecture, white paper, jan 2008.

Dell networking best practices, Dell, 2009.

Dell PowerEdge M1000e Modular Enclosure Architecture, Dell, 2008.

Dynamic Virtual Storage. The Dell Equallogic PS Series, 2010.

Fabric convergence with lossless Ethernet and Fibre Channel over Ethernet (FCoE), Technology Bulletin, Blade Network Technologies, 2008.

Information Storage and Management: Storing, Managing, and protecting Digital Information, EMC Education Services, 2009.

Introdução ao armazenamento unificado com o EMC Celerra, EMC white paper, 2010.

The AMD processor roadmap for industry standard-servers. HP, 2008.

The INTEL processor roadmap for industry standard-servers. HP, 2007.

Veras, Manoel. DATACENTER: Componente Central da Infraestrutura de TI, Brasport, 2009.

PARTE III: SOFTWARE DE VIRTUALIZAÇÃO

8. VMware vSphere

8.1. Introdução

O VMware vSphere é uma solução de VIRTUALIZAÇÃO, fornecida pela empresa VMware, baseada em HYPERVISOR. O vSphere é líder em ambientes corporativos e, segundo a VMware, mais de 170.000 empresas em todo o mundo já utilizam a solução.

O VMware vSphere foi o primeiro sistema operacional lançado para VIRTUALIZAÇÃO já com funcionalidades de CLOUD COMPUTING e utiliza os recursos da VIRTUALIZAÇÃO para transformar DATACENTERS em infraestruturas de computação em nuvem consideravelmente simplificadas, permitindo, segundo a VMware, que as organizações de TI forneçam a próxima geração de serviços de TI flexíveis e confiáveis usando recursos internos e externos com segurança e baixo risco.

A proposta do VMware vSphere é possibilitar reduzir o custo de capital e o custo operacional com a infraestrutura de TI através da VIRTUALIZAÇÃO, além de aumentar o controle sobre o fornecimento de serviços de TI e, enquanto isto, preservar a flexibilidade de escolha entre qualquer tipo de sistema operacional, aplicativo e hardware, hospedados internamente ou que utilizem recursos externos.

O VMware vSphere é uma *suite* para virtualizar aplicações que inclui os HYPERVISORS ESX/ESXi e o software de gerenciamento vCenter Server.

O vSphere permite realizar as seguintes tarefas:

- Rodar múltiplos sistemas operacionais em uma máquina física simultaneamente.
- Otimizar recursos e balancear WORKLOADS entre múltiplas máquinas físicas.

A **Figura 8-1** mostra as relações entre os componentes básicos do vSphere e o vCenter Server.

A versão 4.1 consolida o vSphere como o principal software de VIRTUALIZAÇÃO disponível na atualidade. Na versão 4.1, o ESXi assume o papel de HYPERVISOR principal com a sua integração com o AD da rede Microsoft.

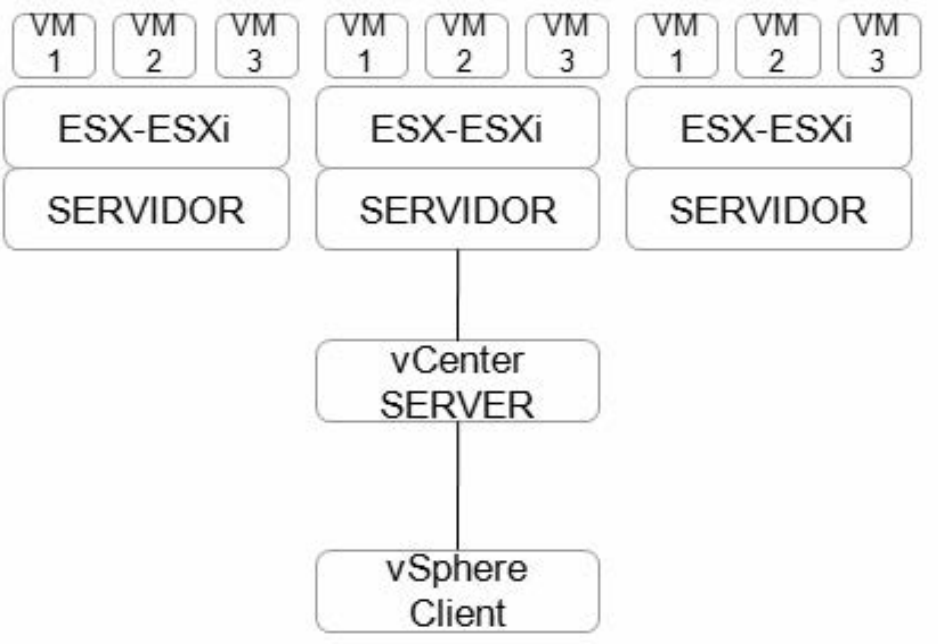


Figura 8-1 vSphere

8.2. Benefícios

O vSphere pode trazer diversos benefícios para a organização. Alguns destes benefícios são listados aqui:

- **Custos reduzidos e melhor eficiência de TI**: o VMware vSphere pode ajudar as organizações a fornecer serviços da TI com mais eficiência, eliminando investimentos de capital desnecessários e reduzindo o custo e a complexidade do gerenciamento e da manutenção da infraestrutura da TI. Segundo a VMware, as organizações que usam o VMware vSphere podem obter taxas de consolidação de cerca de 15:1, com gerenciamento automatizado e alocação dinâmica de recursos a aplicativos em infraestruturas de nuvem interna e/ou externa. O resultado é a quebra do modelo dispendioso de fornecimento de aplicativos e informações ligado a sistemas e arquiteturas específicas.
- **Maior controle de TI por meio da automação do nível de serviço**: na medida em que os negócios passam a depender cada vez mais dos serviços de TI, o fornecimento eficiente de aplicativos pode significar a diferença. Os negócios dependem da área de TI para controlar a qualidade do serviço de fornecimento de aplicativos. O VMware vSphere 4 permite automatizar o fornecimento de SLAs, cobrindo disponibilidade, segurança e escalabilidade, mudando o paradigma do gerenciamento de DATACENTERS baseado em equipamentos para o de fornecimento de serviços. Os novos aplicativos que precisam ser implantados estão protegidos contra a complexidade da infraestrutura de servidores, armazenamento e rede, disponibilizando mais tempo para a equipe de TI se concentrar em demandas atreladas ao aumento do valor dos negócios.

- **Departamentos de TI com possibilidade de escolha:** o VMware vSphere moderniza os ambientes de TI e possibilita fornecer serviços sob demanda, dando liberdade para o departamento de TI escolher o hardware padrão apropriado, a arquitetura de aplicativos, o sistema operacional e decidir sobre a utilização de uma infraestrutura interna ou externa para necessidades comerciais em constante mudança. Segundo a VMware, com o VMware vSphere 4, as organizações mantêm a flexibilidade de escolha, permanecendo independentes de hardware, sistema operacional, aplicativos e provedores de serviços. Isso significa que podem suportar os aplicativos existentes e se sentir confiantes em relação a aplicativos futuros, mantendo, ao mesmo tempo, a flexibilidade para implantar infraestruturas baseadas em nuvem internas e externas.

8.3. Utilização

O vSphere pode ser utilizado da seguinte forma:

- **Consolidar e otimizar o servidor, o armazenamento e o hardware de rede:** O VMware vSphere elimina a proliferação de servidores ao executar aplicativos residentes em máquinas virtuais em um número menor de servidores e com o uso mais eficiente de recursos de armazenamento e rede. As organizações que usam o VMware vSphere podem atingir altas taxas de consolidação por servidor, mediante recursos de gerenciamento de memória e otimização dinâmica. O VMware vSphere reduz a complexidade do gerenciamento de hardware por meio de ampla VIRTUALIZAÇÃO de servidor, armazenamento e hardware de rede. Segundo a VMware, os clientes do VMware vSphere podem reduzir as despesas de capital por aplicativo em pelo menos 50% em média e os custos operacionais (mão de obra) em mais de 60%.
- **Melhorar a continuidade dos negócios por meio de alta disponibilidade e recuperação de desastres simplificados e econômicos:** O VMware vSphere ajuda a criar uma infraestrutura segura que mantém a continuidade dos negócios, independentemente de falhas de hardware ou de paralisações completas do DATACENTER. O VMware vSphere não só elimina o tempo de inatividade de aplicativos resultante de manutenções planejadas de servidor, armazenamento ou de rede, como também oferece alta disponibilidade, em comparação ao tempo de inatividade não planejado, como falhas de servidor. Além disso, o VMware vSphere também simplifica a recuperação de paralisações completas do DATACENTER sem a necessidade de hardwares redundantes e com custos proibitivos.
- **Simplificar as operações de TI:** O VMware vSphere simplifica o gerenciamento operacional de ambientes de teste, desenvolvimento e produção em muitos locais, escritórios remotos e filiais, e inclui todos os tipos de aplicativos e sistemas operacionais. O VMware vSphere facilita o compartilhamento e a substituição de recursos de hardware e simplifica o gerenciamento por meio de conjuntos de políticas comuns, procedimentos operacionais e gerenciamento automatizado em diferentes conjuntos de aplicativos e usuários comerciais.

O VMware vSphere simplifica o provisionamento de serviços e garante um controle uniforme de níveis de serviço, independentemente da infraestrutura física ou da localização real dos serviços. Isso não só reduz a sobrecarga operacional, como também permite a portabilidade de aplicativos em infraestruturas de nuvem interna e/ou externa sem perder os níveis de serviço, nem exigir personalização.

8.4. Componentes da Solução

Uma forma de entender as características do vSphere é entender os serviços propiciados pela plataforma. O vSphere pode ser descrito por meio de serviços de aplicação e serviços de infraestrutura, conforme ilustra a **Figura 8-2**.

- **Serviços de infraestrutura**: conjunto de componentes que virtualizam os recursos de servidor, armazenamento e rede, agregam e alocam esses recursos sob demanda a aplicativos com base nas prioridades dos negócios.
- **Serviços de aplicativos**: conjunto de componentes que oferecem controles incorporados no nível de serviço para todos os aplicativos executados no VMware vSphere, independentemente do tipo de aplicativo ou sistema operacional.

O VMware vCenter Server é a ferramenta para administração dos serviços de aplicativos e infraestrutura e também para automação de tarefas operacionais com grande visibilidade de todos os aspectos do ambiente VMware vSphere.

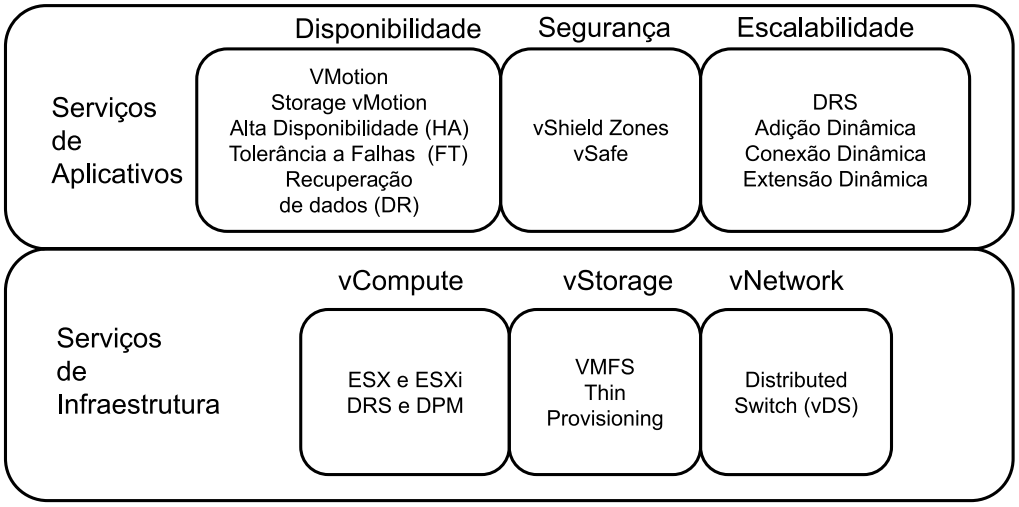


Figura 8-2 Componentes do vSphere 4.0

8.4.1. Serviços de Infraestrutura

- **VMware vCompute**: O VMware vCompute trata de serviços de infraestrutura que virtualizam recursos de servidor com eficiência e os incorporam em *pools* lógicos que podem ser alocados com precisão a aplicativos.
  - O VMware ESX e o VMware ESXi oferecem uma camada de virtualização robusta, de bom desempenho, que abstrai os recursos de hardware do servidor e

permite o compartilhamento desses recursos entre várias máquinas virtuais. Os recursos de gerenciamento de memória e programação do VMware ESX e do ESXi oferecem altas taxas de consolidação e bom desempenho de aplicativo.

- O VMware Distributed Resource Scheduler (DRS) incorpora recursos de computação em vários clusters e os aloca dinamicamente a máquinas virtuais com base nas prioridades dos negócios, reduzindo a complexidade do gerenciamento por meio de automação.
  - O VMware Distributed Power Management (DPM), incluído no VMware DRS, automatiza o uso eficiente de energia nos clusters do VMware DRS por meio da otimização contínua do consumo de energia do servidor dentro de cada cluster.

- **VMware vStorage:** O VMware vStorage trata de serviços de infraestrutura que abstraem recursos de armazenamento da complexidade dos sistemas subjacentes de hardware, para proporcionar o uso mais eficiente da capacidade de armazenamento em ambientes virtualizados.
  - O VMware vStorage Virtual Machine File System (VMFS) é o sistema de arquivos de cluster de alto desempenho do vSphere que possibilita um compartilhamento eficiente e controla o acesso simultâneo de servidores virtualizados ao armazenamento. O vSphere permite a expansão dinâmica dos volumes VMFS. Se a LUN aumenta de tamanho, o VMFS volume grow permite estender o volume e aumentar o tamanho dinamicamente.
  - O VMware vStorage Thin Provisioning oferece alocação dinâmica de capacidade de armazenamento, permitindo que compras de armazenamento sejam adiadas até que se tornem realmente necessárias, reduzindo os gastos com armazenamento em até 50%, segundo a VMware. O VMware vStorage Thin Provisioning não é equivalente ao Thin Provisioning do storage, já descrito na capítulo anterior.

O VMware vSphere introduz o *Pluggable Storage Architecture* (PSA), que permite que parceiros de storage escrevam plug-ins para capacidades específicas. O *Native Multipath Driver* (NMP) fornecido pela VMware é utilizado por default, mas se o fornecedor de storage oferece um driver específico, este pode ser utilizado para gerenciar a interface entre o ESX e o storage.

O VMware Paravirtualized SCSI (PVSCSI) é um driver de uso geral para adaptadores de alta performance que oferece maior *throughput* e menor utilização de CPU para máquinas virtuais. São utilizados em ambientes com aplicações intensivas em I/O.

O VMware VMDirectPath I/O permite que VMs com aplicações que acessam intensamente e frequentemente o I/O possam acessar diretamente o hardware, melhorando o desempenho. Esta funcionalidade mapeia uma HBA para uma VM. É importante ressaltar que funcionalidades como VMotion não podem ser utilizadas neste caso.

- **VMware vNetwork**: O VMware vNetwork trata de serviços de infraestrutura que proporcionam administração e gerenciamento de rede em ambientes virtuais e inclui:
  - O VMware vNetwork Distributed Switch (vDS) simplifica e melhora o provisionamento, a administração e o controle da rede de máquinas virtuais em ambientes VMware vSphere em relação à solução standard vSwitch anteriormente fornecida pela versão do VMware. Além disso, permite que switches virtuais distribuídos por terceiros, como o Cisco Nexus 1000v, sejam usados em ambientes VMware vSphere, oferecendo aos administradores de rede interfaces familiares para controle da qualidade do serviço no nível da máquina virtual.

  O vDS introduz outros benefícios como:

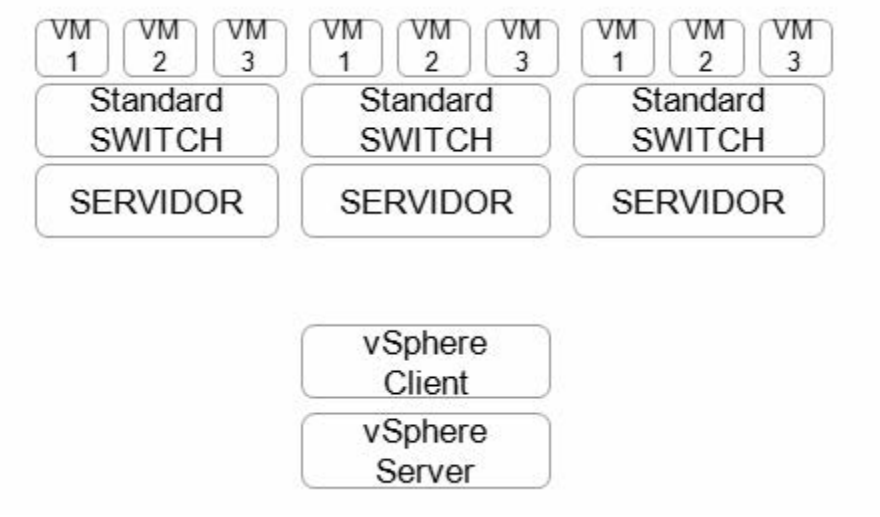
  - VLANs privadas que permitem aos usuários restringirem a comunicação entre VMs na mesma VLAN ou em um segmento de rede, reduzindo o número de subnets necessárias para certas configurações.
  - Network VMotion, que é um acompanhamento do estado da máquina virtual quando uma VM se move de um host para outro.
  - Suporte a switches virtuais de terceiros como a Nexus 1000v, além de possibilitar fazer o *shaping* do tráfego bidirecional (da VM para a rede e da rede para a VM).

O vSphere introduz o VMXNET3, que é uma NIC paravirtual para sistemas operacionais convidados que permite VLAN off-loading e TCP segmentation offloading (TSO) over IPv6.

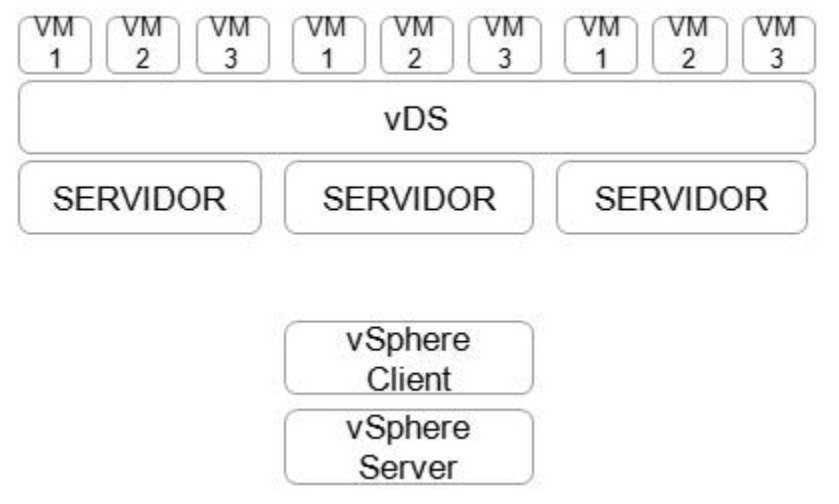
O suporte ao IPV6 foi melhorado, permitindo que o suporte seja estendido ao *vmkernel* e ao service console.

O VMDirectPath PCI permite que dispositivos PCI sejam vinculados diretamente a uma VM para melhoria de desempenho. As máquinas virtuais específicas que utilizam o VMDirectPath PCI perdem a chance de utilizar funcionalidades como o VMotion.

A **Figura 8-3** ilustra o avanço da solução de VIRTUALIZAÇÃO para a rede. Na **Figura 8-3 (a)** os switches são configurados e gerenciados individualmente (*standard switch*). Na **Figura 8-3 (b)**, o gerenciamento dos switches (*distributed switch*) é independente do número de hosts.



(a) standard switch



(b) vDS

Figura 8-3 Evolução do Switch Virtual no VMware

8.4.2. Serviços de Aplicativos

Os serviços de aplicativos do VMware vSphere oferecem controles incorporados sobre níveis de serviços de aplicativos, como disponibilidade, segurança e escalabilidade, e podem ser habilitados de maneira simples e uniforme em qualquer aplicativo executado em máquinas virtuais VMware.

- **Disponibilidade:** os serviços de disponibilidade permitem que o setor de TI forneça aplicativos com níveis variados de alta disponibilidade, de acordo com a prioridade e a necessidade, sem precisar de hardwares complexos redundantes nem software de cluster.

- O VMware vMotion elimina a necessidade de programar o tempo de inatividade de aplicativos, em virtude da manutenção de servidor planejada por meio da migração ao vivo de máquinas virtuais entre servidores, sem interrupção das atividades dos usuários nem perda de serviço.
- O VMware Storage vMotion elimina a necessidade de programar o tempo de inatividade de aplicativos, em virtude da manutenção de servidor planejada ou durante as migrações de armazenamento por meio da migração ao vivo de discos de máquinas virtuais, sem interrupção das atividades dos usuários nem perda de serviço.
- O VMware High Availability (HA) permite fazer a reinicialização automática e econômica em minutos para todos os aplicativos em caso de falhas de hardware ou do sistema operacional. O VMware HA permite obter HA por servidor físico.
- O VMware Fault Tolerance (FT) permite obter disponibilidade contínua para qualquer aplicativo sem perda de dados nem tempo de inatividade, em caso de falhas de hardware. O VMware FT permite obter alta disponibilidade por máquina virtual.
- O VMware Data Recovery (DR) permite fazer backup e recuperação simples e econômicos, sem agente, de máquinas virtuais. O VMware DR pode ser integrado a soluções comerciais de backup e é baseado no VSS da Microsoft. Importante salientar que o VDR não é o VCB.

- **Segurança:** os serviços de segurança permitem que a TI forneça aplicativos com o nível apropriado de segurança de forma operacionalmente eficiente.
  - O VMware vShield Zones simplifica a segurança de aplicativos ao aplicar políticas de segurança corporativa no nível do aplicativo em um ambiente compartilhado mantendo, ao mesmo tempo, a confiança e a segmentação de rede de usuários e dados confidenciais.
  - O VMware VMsafe permite o uso de produtos de segurança que funcionam em conjunto com a camada de virtualização, para oferecer a máquinas virtuais níveis mais altos de segurança do que os oferecidos por servidores físicos.
- **Escalabilidade:** os serviços de escalabilidade permitem a TI fornecer o volume adequado de recursos a cada aplicativo, com base nas necessidades, sem interrupções.
  - O VMware DRS balanceia dinamicamente a carga de recursos do servidor para oferecer o recurso certo ao aplicativo certo com base na prioridade dos negócios, permitindo que o aplicativo cresça ou diminua, conforme necessário.
  - A Adição Dinâmica permite a adição de CPU e memória a máquinas virtuais quando necessário, sem interrupções nem tempo de inatividade.
  - A Conexão Dinâmica permite a adição ou a remoção de armazenamento virtual e dispositivos de rede para ou de máquinas virtuais sem interrupções nem tempo de inatividade.

- A Extensão Dinâmica de discos virtuais permite a adição de armazenamento virtual para a execução de máquinas virtuais sem interrupções nem tempo de inatividade.

- **vApp:** o vSphere suporta a vApp, entidade lógica formada por uma ou mais máquinas virtuais que utilizam o OPEN VIRTUALIZATION FORMAT padrão do setor. O vApp permite especificar e encapsular componentes de um aplicativo de vários níveis.

8.5. Novidades com o vSphere 4.1

8.5.1. Introdução

Entre as principais novidades da versão 4.1, pode-se citar:

- **Afinidade de HOST no DRS:** permite estabelecer uma política granular para o movimento de máquinas virtuais como, por exemplo, restringir o movimento de uma máquina virtual para um host específico devido ao impacto no licenciamento.
- **Compressão de Memória:** melhoria de até 30% no desempenho das aplicações com a redução da contenção de memória.
- **Relatórios de desempenho:** estabelece diversas estatísticas para o desempenho do storage.
- **Integração de Array:** fornece novas APIs (VMware vStorage APIs for Array Integration – VAAI) para a interface entre VMs e arrays de storage.
- **Segurança:** Integração com AD permite a autenticação no host vSphere, possibilitando o gerenciamento centralizado.
- **Arquitetura interoperável e aberta:** fornece novas APIS para integração do array, melhorando o desempenho e a disponibilidade.
- **Aumento do Suporte:** lista HCL expandida.

O VMware vSphere 4.1 introduz uma série de melhorias em relação ao vSphere 4.0.

8.5.2. Storage I/O Control (SIOC)

O SIOC permite estabelecer prioridades de qualidade de serviço de I/O para cada VM dentro de um pool de servidores ESX, garantindo o acesso aos recursos de armazenamento de forma a atender os níveis de serviço adequados para cada aplicação.

O que o SIOC faz é fornecer mais slots de I/O para máquinas virtuais de aplicações críticas. A **Figura 8-4** ilustra o funcionamento do Storage I/O Control. A aplicação 1 tem prioridade no uso dos recursos do storage sobre as aplicações 2 e 3.

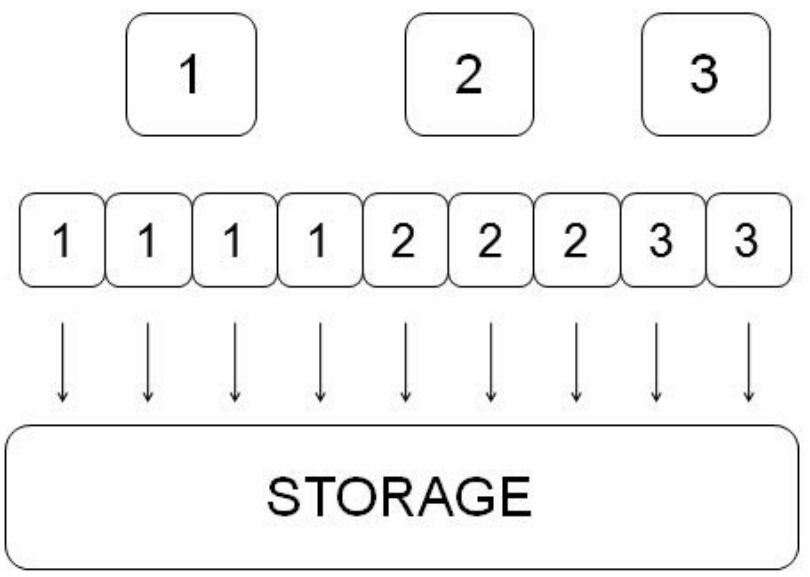


Figura 8-4 Storage I/O Control

8.5.3. Network I/O Control (NIOC)

O NIOC é uma nova funcionalidade de gerenciamento de tráfego do vDS. Permite estabelecer prioridades de qualidade de serviço para a rede por tipo de fluxo, garantindo o acesso aos recursos.

O NIOC permite priorizar e identificar os seguinte tráfegos que deixam os hosts ESX/ESXi por um uplink vDS:

- Tráfego VM.
- Tráfego de gerenciamento.
- iSCSI.
- NFS.
- FT logging.
- VMotion.

A **Figura 8-5** ilustra o funcionamento do NIOC com priorização do tráfego. O NIOC se aplica no tráfego que deixa o host ESX-ESXi . O NIOC se aplica particularmente a tráfegos que convergem em interfaces de 10GbE. Cada um dos tráfegos receberá um nível de serviço em função da prioridade estabelecida. Os schedulers priorizam o tráfego de acordo com os *shares* estabelecidos no vDS.

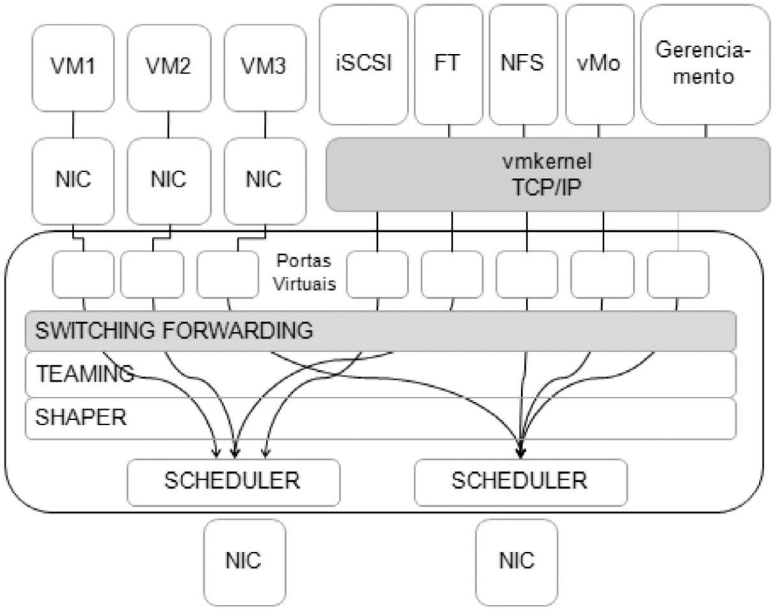


Figura 8-5 Network I/O Control

8.5.4. vSphere vCenter Server 4.1

As principais melhorias introduzidas no vSphere 4.1 vCenter foram:

- Transição para 64 bits. Siginifica que é necessário agora um SO de 64 (Windows Server 2008 R2) bits para suportar o vCenter Server. As limitações de endereçamento de memória (4GB) para a versão 32 bits acabam.
- Melhorias na escalabilidade e desempenho. Decorrente da versão de 64 bits. Melhora no startup e no login do cliente. Melhor integração entre HA/FT e DRS/DPM.

Os novos limites para o vSphere são mostrados na **Tabela 8-1**.

Tabela 8-1 Comparação vSphere 4.0 com vSphere 4.1

Limites	vSphere 4	vSphere 4.1	Melhoria
VMs por host	320	320	-
Hosts por cluster	32	32	-
VMs por cluster	1280	3000	3x
Hosts por vCenter Server	300	1000	3x
vSphere clients concorrentes	30	120	4x
Hosts por DATACENTER	100	500	5X

VMs por DATACENTER	2500	5000	2X
Linked Mode	10000	30000	3x
vCenter Server operações concorrentes	100	500	5x
Distributed Switches por vCenter Server	16	32	2x
Hosts por distributed switches	64	350	5x
vMotion sobre 10G (rede)	2	8	4x
vMotion sobre FC, iSCSI, NFS (datastore)	8	128	16x
Storage vMotion sobre FC, iSCSI, NFS (datastore)	8	8	1x
vMotion (HOST)	2	8	4x
Storage vMotion (HOST)	2	2	1x

- Suporte a integração do ESX-ESXi com o Microsoft AD. O usuário pode ser autenticado diretamente pelo AD nos servidores ESX/ESXi.
- Melhorias no VMware vCenter Host Profiles, que permitem a configuração ou reconfiguração de vários hosts ao mesmo tempo, agora integrados com o AD.
- Melhorias no VMware vCenter Orchestrator em termos de escalabilidade e desempenho com a ida para 64 bits e agora a opção do cliente em 32 ou 64 bits.
- Melhoria no VMware vCenter Update Manager com a melhoria da escalabilidade e desempenho.
- Melhorias no VMware vCenter Converter 3 que agora suporta a conversão virtual-to-virtual (V2V). Permite importar facilmente máquinas virtuais baseadas em HYPER-V.

8.5.5. Configurações Máximas do vSphere 4.1

As tabelas a seguir ilustram as configurações máximas para Máquina Virtual, ESX e vCenter Server para o vSphere 4.1.

Tabela 8-2 Máquina Virtual

CPU	
CPUs virtuais por VM	8
Memória	
RAM por VM	255 GB
Tamanho do swap file da VM	255 GB
Storage	
Adaptadores Virtuais SCSI por VM	4
Tamanho de disco	2 TB
Controladora IDE por VM	1
Dispositivos IDE por VM	4
Rede	
NICs virtuais por VM	10
Portas Periféricas	
Controladores USB por VM	1
Dispositivos USB conectados a VM	10
Portas Paralelas por VM	3
Portas Seriais	4

Tabela 8-3 ESX

CPU	
CPU lógica por Host	128
VM por host	320
CPU virtual por host	512
CPU virtual por core	25
Discos Virtuais	16
Memória	
RAM por Host	1 TB
RAM máxima por Service Console	800 MB
Número de swap files	1 por VM
Storage	
iSCSI	
LUNS por HOST	256
QLogic ou Broadcom 1GB iSCSI initiator por HOST	4
Broadcom 10Gb iSCSI initiator por HOST	4
Broadcom 1Gb-10Gb iSCSI target por HBA	64
iSCSi target por software	256
NAS	
NFS por host	64
Fibre Channel	
LUNs por Host	256
Tamanho da LUN	2 TB

LUN ID	255
Número de paths para LUN	32
Número total de paths por Servidor	1024
Número de HBAs	8
Portas HBA	16
Targets por HBA	256
VMFS	
Tamanho do RDM	2 TB
Tamanho do Volume	64 TB
Volumes por Host	256
Hosts por Volume	64
Rede	
Portas Ethernet de 1GB Intel PCI-X	32
Portas Ethernet de 1GB Intel PCI-E	24
Portas Ethernet de 10Gb Neterion ou NetXen ou Broadcom	4
Portas Virtuais vDS e vSS por Host	4096
Máximo de portas ativas vDS e vSS por Host	1016
Swtiches distribuídos por vCenter	32
Swtiches distribuídos por Host	16
Host por Switch Distribuído	350
Resource Pool e Cluster	
Host por Cluster	32
VM por Cluster	3000
VM por host	320

Resource Pool por Cluster	512
Resource Pool por Host	4096
Children por Resource Pool	1024

Tabela 8-4 vCenter

Hosts por VCenter	1000
VM por vCenter Server	10.000
VM por vCenter (registradas)	15.000
vCenters Servers Linkados	10
Hosts em vCenter Servers Linkados	3000
VM em vCenter Linkados	30.000
VM em vCenter Linkados (registradas)	50.000
Operações de provisionamento concorrente por Host	4
Operações de provisionamento concorrente por datastore	4
Operações concorrentes com VMotion por Host (1GB/s)	4
Operações concorrentes com VMotion por Host (10GB/s)	8
Operações concorrentes com storage VMotion por Host	2
Operações concorrentes com storage VMotion por Datastore	8
Hosts Scan em um vCenter Server	1000
VM Scan em um vCenter Server	10.000

8.6. Edições

As edições do vSphere 4.1 são descritas na **Tabela 8-5,** de acordo com as características citadas abaixo.

8.6.1. Características

1. Consolidação: Converte sistemas físicos em máquinas virtuais e possibilita a migração viva.
2. Disponibilidade: fornece alta disponibilidade e tolerância a falhas para as aplicações.
3. Gerenciamento Automatizado de Recursos: disponibiliza balanceamento de carga, gerenciamento de energia e migração viva do storage sem intervenção manual.
4. Operação Simplificada: operação de rede avançada e templates para configuração de hosts.

8.6.2. Kits e Edições

- **VMware vSphere Essentials**: projetado especialmente para ambientes de TI com menos de vinte servidores físicos, estes kits entregam VIRTUALIZAÇÃO de classe empresarial com gerenciamento integrado para pequenos negócios.
- **VMware vSphere Essentials PLUS**: projetado especialmente para ambientes de TI com menos de vinte servidores físicos, este kit entrega VIRTUALIZAÇÃO de classe empresarial com gerenciamento integrado e funcionalidades de continuidade de negócios para pequenos negócios
- **VMware vSphere Essentials for Retail & Branch Offices**: kits projetados para atender a médias e grandes empresas que querem virtualizar sites remotos. Propiciam soluções turnkey para maior agilidade dos negócios e continuidade em todos os locais remotos.
- **VMware vSphere Essentials PLUS for Retail & Branch Offices**: em relação à edição anterior, permite obter também recursos de disponibilidade.
- **VMware vSphere Standard Edition**: edição de nível de entrada para consolidação básica de aplicações, reduz o custo de hardware e acelera o desenvolvimento de aplicações.
- **VMware vSphere Advanced Edition**: esta edição é uma opção estratégica para consolidação de todas as aplicações contra o downtime planejado e não planejado e assim aumentar a disponibilidade.
- **VMware vSphere Enterprise Edition**: esta edição inclui todas as funcionalidades necessárias para transformar os DATACENTERS e simplificar o ambiente de CLOUD COMPUTING, possibilitando a entrega melhorada de serviços de TI.
- **VMware Enterprise Plus Edition**: esta edição inclui todas as funcionalidades da versão Enterprise e ainda as opções de simplificação da operação de rede, além de fornecer templates que simplificam a instalação.

Tabela 8-5 Kits e Edições do vSphere

	A	B	C	D	E	F	G	H
1	X	X	X	X	X	X	X	X
2		X		X		X	X	X
3							X	X
4								X

Existem também os VMware vSphere Acceleration kits:

- vSphere Advanced Acceleration kit for 6 processors.
- vSphere Midsize Acceleration kit for 6 processors.
- vSphere Enterprise Plus Acceleration kit for 8 processors.

8.6.3. Add-ons e Produtos Adicionais

- **VMware vCenter Server**: permite obter gerenciamento centralizado para todo o ambiente virtual e disponibiliza as funcionalidades de migração viva. Pode gerenciar centenas de máquinas virtuais em diversas localizações, facilitando a administração e o rápido provisionamento de máquinas virtuais. Importante ressaltar que o vCenter Server é licenciado separadamente e por instância. Os clientes já possuem o vCenter Server quando adquirem as edições Essentials. O VMware vCenter Server é comercializado de duas formas: Foundation e Standard.
- **VMware Data Recovery**: é disponível para aquisição como módulo add-on para a versão standard. O add-on é tratado com um upgrade da licença do VMware Standard.
- **Cisco Nexus 1000V**: É um switch por software integrado à edição do vSphere Enterprise Plus que disponibiliza novos serviços de rede.

A ferramenta da VMware vSphere Purchase Advisor, encontrada em www.vmware.com, cuja tela inicial é mostrada na **Figura 8-6**, facilita a definição da escolha da edição a ser utilizada pela organização que está adquirindo o vSphere.

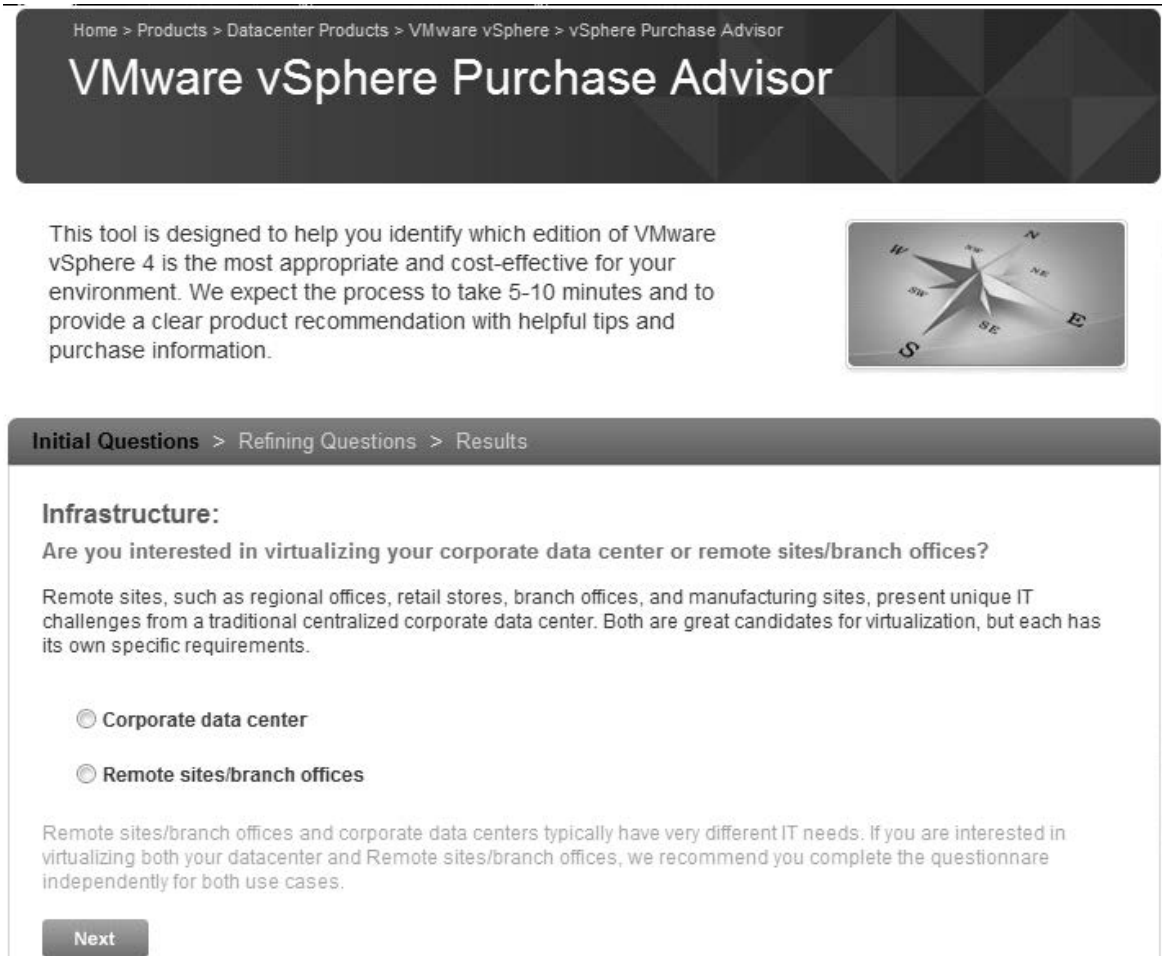

Home > Products > Datacenter Products > VMware vSphere > vSphere Purchase Advisor
VMware vSphere Purchase Advisor
This tool is designed to help you identify which edition of VMware vSphere 4 is the most appropriate and cost-effective for your environment. We expect the process to take 5-10 minutes and to provide a clear product recommendation with helpful tips and purchase information.
Initial Questions > Refining Questions > Results
Infrastructure:
Are you interested in virtualizing your corporate data center or remote sites/branch offices?
Remote sites, such as regional offices, retail stores, branch offices, and manufacturing sites, present unique IT challenges from a traditional centralized corporate data center. Both are great candidates for virtualization, but each has its own specific requirements.
Corporate data center
Remote sites/branch offices
Remote sites/branch offices and corporate data centers typically have very different IT needs. If you are interested in virtualizing both your datacenter and Remote sites/branch offices, we recommend you complete the questionnare independently for both use cases.
Next

Figura 8-6 VMware Purchase Advisor

8.7. ESX e ESXi

8.7.1. Introdução

A solução de VIRTUALIZAÇÃO da VMware é baseada nos HYPERVISORS ESX/ESXi, cuja arquitetura é mostrada na **Figura 8-7**. A ideia básica é abstrair a carga de trabalho do hardware e criar um pool de recursos que podem ser alocados dinamicamente em resposta às mudanças nas condições do negócio. Com isto, é possível carregar vários aplicativos em um único servidor físico sem haver possibilidade de que um mau funcionamento de um aplicativo possa interferir no outro. Considerando que os servidores em geral estão subutilizados em instalações que não utilizam a VIRTUALIZAÇÃO, a VIRTUALIZAÇÃO com o ESX/ESXi possibilita otimizar o uso dos recursos de TI.

A VIRTUALIZAÇÃO avançou e trouxe o desenvolvimento de outras funcionalidades. Na prática o servidor ESX-ESXi é o núcleo da VIRTUALIZAÇÃO, mas o vSphere é muito mais do que o ESX-ESXi. As funcionalidades oferecidas justificam melhor o projeto de VIRTUALIZAÇÃO aumentando, inclusive, o nível de disponibilidade do novo ambiente.

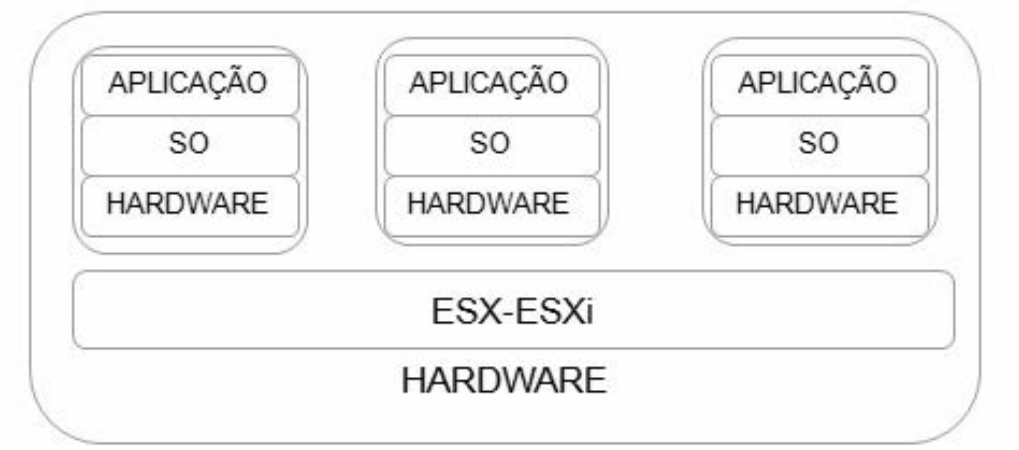


Figura 8-7 Arquitetura do VMware ESX-ESXi

A **Figura 8-8** ilustra o posicionamento do ESX-ESXi.

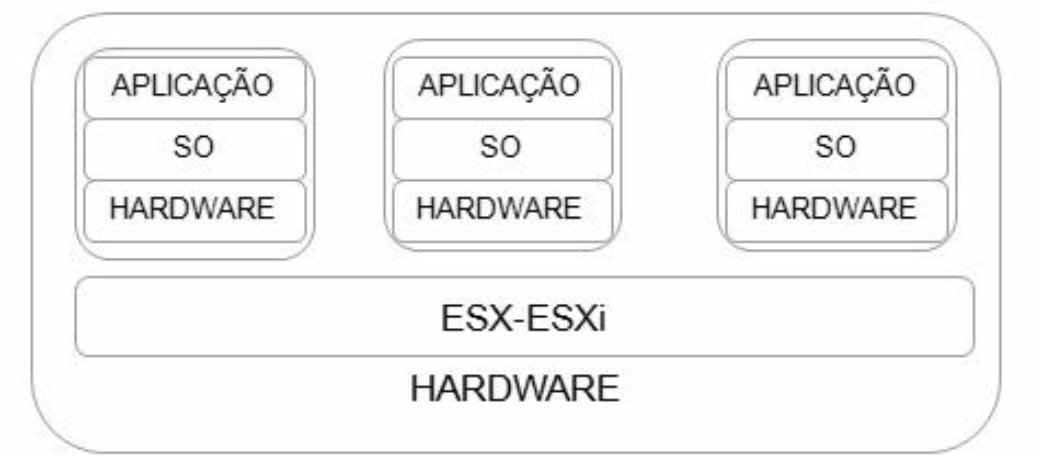


Figura 8-8 ESX e ESXi

O VMware ESX e o VMware ESXi podem ser implantados como parte da plataforma vSphere ou do conjunto de produtos VMware View para permitir gerenciamento centralizado e melhor qualidade dos serviços para os aplicativos de DATACENTER e *desktops* corporativos, permitindo que gerentes de TI:

- Implementem a consolidação e a contenção de servidores de produção.
- Forneceçam proteção avançada para a continuidade dos negócios por um custo reduzido.
- Gerenciem e controlem desktops virtuais centralizados.
- Simplifiquem o desenvolvimento e o teste de software.
- Hospedem novamente aplicativos legados.

O VMware ESXi pode ser implantado como solução de VIRTUALIZAÇÃO de servidor único. O VMware vSphere client, disponível gratuitamente, permite que o ESXi crie e gerencie máquinas virtuais. Como o ESXi é funcionalmente equivalente ao ESX, é possível fazer um *upgrade* do ESX para o ESXi . O ESXi permite a integração com o Microsoft AD.

O VMware ESXi é a arquitetura de HYPERVISOR mais recente da VMware. Ele possui uma arquitetura ultrafina que não depende de um sistema operacional de uso geral e ainda oferece a mesma funcionalidade e desempenho do VMware ESX. O VMware ESXi cria um novo patamar para a segurança e a confiabilidade, pois sua base de código menor representa uma "superfície de ataque" reduzida, com menos código para corrigir. Com o pequeno espaço ocupado e a mesma confiabilidade de um hardware, o VMware ESXi pode ser integrado diretamente aos servidores x86 padrão do setor pelos fabricantes líderes de servidores, como Dell, IBM e HP. O VMware ESXi foi desenvolvido objetivando a simplicidade. A inicialização por menu e as configurações automáticas tornam o ESXi a maneira mais fácil de conhecer a VIRTUALIZAÇÃO da VMware.

Para instalar o ESXi 4.1 é necessário:

- Plataforma de Servidor suportada de 64 bits (ver VMware systems compatibility guide).
- Memória RAM de 2Gb RAM no mínimo.
- Uma ou mais NICs de 1Gb ou 10Gb.
- Um ou mais controladores de disco incluindo Dell PERC, HP Smart Array ou IBM ServerRAID.
- Discos SAS ou Discos SATA com controladora SATA ou SAS.

8.7.2. Funcionamento

O VMware ESX e o VMware ESXi são instalados diretamente no hardware do servidor, inserindo uma camada de VIRTUALIZAÇÃO entre o hardware e o sistema operacional.

O VMware ESX e o ESXi particionam um servidor físico em várias máquinas virtuais seguras e portáveis que podem ser executadas, lado a lado, no mesmo servidor físico. Cada máquina virtual representa um servidor completo – com processadores, memória, rede, armazenamento e BIOS – de modo que o sistema operacional e os aplicativos de software possam ser instalados e executados na máquina virtual sem qualquer modificação.

As máquinas virtuais também são totalmente isoladas umas das outras pela camada de VIRTUALIZAÇÃO, evitando, desta maneira, que uma falha ou um erro de configuração em uma máquina virtual afete as demais.

Compartilhar os recursos do servidor físico entre várias máquinas virtuais aumenta a utilização do hardware e reduz consideravelmente os custos de capital. A arquitetura "bare metal" permite ao VMware ESX e ao ESXi controle total sobre os recursos de ser-

vidor alocados a cada máquina virtual, além de oferecer desempenho e escalabilidade às máquinas virtuais quase idênticos aos da original.

O VMware ESX e o ESXi fornecem às máquinas virtuais recursos incorporados de alta disponibilidade, gerenciamento de recursos e funções de segurança que proporcionam melhores níveis de serviço para os aplicativos de software do que os ambientes físicos estáticos.

8.7.3. Diferenças

O VMware ESX e o VMware ESXi são HYPERVISORS bare metal instalados diretamente no hardware de servidor. Ambos fornecem desempenho e escalabilidade; a diferença está na arquitetura e no gerenciamento operacional do VMware ESXi. O VMware ESX conta com um sistema operacional Linux, chamado de console de serviço, para desempenhar algumas funções de gerenciamento, inclusive a execução de *scripts* e a instalação de agentes de terceiros para monitoramento de hardware, backup ou gerenciamento de sistemas.

O console de serviço foi removido do VMware ESXi, reduzindo significativamente a ocupação de espaço. Ao remover o console de serviço, o VMware ESXi segue a tendência de migrar o recurso de gerenciamento da interface local de linha de comando para ferramentas de gerenciamento remoto. A função do console de serviço é substituída pelas interfaces de linha de comando remotas e adere aos padrões de gerenciamento do sistema.

8.7.4. Recursos

Os recursos gerais do ESX-ESXi são descritos aqui:

- **Arquitetura de 64 bits.** Apresenta desempenho e suporte aprimorados para até 1 TB de RAM em hosts físicos.
- **Otimizações de desempenho para cargas de trabalho virtualizadas.** O VMware ESX e o ESXi passaram por otimizações de desempenho para aplicativos específicos essenciais aos negócios, como o banco de dados Oracle, banco de dados Microsoft SQL Server e o correio eletrônico Microsoft Exchange. Permite obter até 8.900 transações de bancos de dados por segundo, 200.000 operações de I/O por segundo e até 16.000 caixas de correio Exchange por host.
- **Desempenho aprimorado para armazenamento iSCSI.** Permite utilizar uma combinação entre os novos drivers SCSI otimizados para VIRTUALIZAÇÃO e as otimizações de pilha de armazenamento no nível do VMkernel para aprimorar significativamente o desempenho dos aplicativos de I/O intenso, como bancos de dados e aplicativos de mensagens.
- **Suporte a máquinas virtuais maiores e hardware de servidor potente.** Aproveita os sistemas de hardware com até 64 núcleos físicos de CPU, 256 CPUs virtuais, 1 TB de RAM e centenas de máquinas virtuais em um único host para facilitar a consolidação em larga escala e os projetos de recuperação de desastres. Permite configurar máquinas virtuais com até 255 GB de RAM.

- **Suporte a SMP virtual de oito vias.** O VMware Virtual Symmetric Multi-Processing (SMP) melhora o desempenho das máquinas virtuais ao permitir que uma única máquina virtual utilize, de maneira simultânea, até oito processadores físicos. O VMware Virtual SMP permite a VIRTUALIZAÇÃO dos aplicativos corporativos que mais utilizam a CPU, como bancos de dados, ERP e CRM.
- **VMware VMsafe.** O VMware VMsafe é uma nova tecnologia de segurança que ajuda a proteger cargas de trabalho virtualizadas de uma maneira que antes não era possível com as máquinas físicas. O VMsafe fornece um conjunto de APIs de segurança que permite aos produtos de segurança de terceiros obter a mesma visibilidade do VMware ESX ou do ESXi durante a operação de uma máquina virtual para identificar e eliminar malware. Essa proteção avançada é obtida por meio da visibilidade granular dos recursos de hardware da máquina virtual, como memória, CPU e disco, e de seus sistemas de I/O.
- **VMDirectPath para máquinas virtuais.** Aumenta a eficiência da CPU para aplicativos que exigem acesso frequente aos dispositivos de I/O ao permitir que máquinas virtuais selecionadas acessem diretamente os dispositivos de hardware subjacentes.
- **Melhor gerenciamento de energia.** Melhora o aproveitamento de energia com voltagem dinâmica e escalabilidade de frequência e suporte para Intel SpeedStep e AMD PowerNow!.

8.7.5. Arquitetura

Descrevem-se aqui os principais recursos da arquitetura.

- **Arquitetura de HYPERVISOR bare metal de 64 bits.** Permite obter desempenho, confiabilidade e escalabilidade quase idênticas à da máquina original com a tecnologia de HYPERVISOR executada diretamente no hardware de servidor, sem a necessidade de um sistema operacional de host.
- **Arquivos de disco virtual.** Permite utilizar os arquivos VMDK (*virtual machine disk*, disco de máquina virtual) para fornecer às máquinas virtuais acesso a seus próprios datastores privados e, aos administradores de TI, a flexibilidade de criar, gerenciar e migrar o armazenamento de máquinas virtuais como arquivos independentes que podem residir em equipamentos de armazenamento compartilhado.
- **VMware vStorage VMFS.** Permite eliminar pontos únicos de falha e equilibrar os recursos de armazenamento ao implementar o armazenamento compartilhado para máquinas virtuais com o VMware vStorage Virtual Machine File System (VMFS), um sistema de arquivos clusterizado que permite que vários hosts do VMware ESX acessem um único arquivo VMDK ao mesmo tempo. O VMFS é suportado por arrays de armazenamento SAN Fibre Channel, SAN iSCSI e NAS de modo transparente para os donos e usuários de aplicativos.
- **Inicialização a partir de SAN.** Permite eliminar a necessidade de fazer backup separado de discos locais de servidor conectados ao executar os hosts do VMware ESX em configurações sem disco de servidores.

- **Rede virtual.** Permite aos clientes criar redes complexas entre as máquinas virtuais que residem em um único host ou entre várias instalações de hosts do VMware ESX e do ESXi para implantações de produção ou para desenvolvimento e teste. Configura cada máquina virtual com um ou mais NICs virtuais, cada um deles com seu próprio endereço MAC e IP, para que as máquinas virtuais não sejam diferenciadas das máquinas físicas. Cria uma rede simulada em um host do VMware ESX com switches virtuais conectados a máquinas virtuais. Permite utilizar LANs virtuais (VLANs) para sobrepor uma LAN lógica a LANs físicas, a fim de isolar o tráfego de rede para fins de segurança e separação de carga. Permite modificar as configurações de rede sem precisar alterar fisicamente o cabeamento e as configurações de switch atuais.

8.7.6. Gerenciamento

O VMware ESX gerencia recursos para aprimorar o desempenho e aumentar as taxas de consolidação. Os principais recursos de gerenciamento são:

- **Gerenciamento de recursos para máquinas virtuais.** Permite definir políticas avançadas de alocação de recursos para máquinas virtuais, a fim de aumentar os níveis de serviço para aplicativos de software. Permite estabelecer cotas mínimas, máximas e proporcionais de CPU, memória, disco e largura de banda da rede. Permite modificar as alocações enquanto as máquinas virtuais estão em execução.
- **Virtualização inteligente da CPU.** Permite gerenciar como os processos das máquinas virtuais são executados com a programação inteligente de processos e balanceamento de carga em todas as CPUs disponíveis no host físico.
- **Comprometimento da RAM acima do limite.** Permite aumentar a utilização da memória configurando a memória da máquina virtual para ultrapassar a memória do servidor físico, permitindo que um número maior de máquinas virtuais sejam executadas em um host do VMware ESX ou do ESXi.
- **Compartilhamento transparente de página (eliminação de duplicação de memória).** Permite utilizar a memória RAM física de maneira mais eficiente, armazenando páginas de memória de forma idêntica em várias máquinas virtuais, uma única vez.
- **Aumento da capacidade da memória.** Permite mudar a memória RAM dinamicamente de máquinas virtuais ociosas para cargas de trabalho ativas. O aumento da capacidade da memória induz artificialmente uma pressão sobre a memória dentro de máquinas virtuais ociosas, obrigando-as a usar suas próprias áreas de paginação e liberar memória para as máquinas virtuais ativas.
- **Modelagem de tráfego de rede.** Garante que máquinas virtuais essenciais tenham acesso prioritário à largura de banda da rede. O tráfego de rede das máquinas virtuais pode ser priorizado com base em "cotas justas". O Network Traffic Shaper gerencia o tráfego de rede das máquinas virtuais para adaptar-se a restrições de largura de banda (pico, utilização média e sobrecarga).
- **Priorização do tráfego de I/O de armazenamento**. Permite garantir que máquinas virtuais essenciais tenham acesso prioritário a dispositivos de armazenamento pela priorização do tráfego de I/O com base em "cotas justas".

- **Melhor gerenciamento de energia.** Permite melhorar o aproveitamento de energia com voltagem dinâmica e escalabilidade de frequência e suporte para Intel SpeedStep e AMD PowerNow.

Interfaces de Gerenciamento

Várias interfaces de gerenciamento estão disponíveis para gerenciar com mais eficiência os ambientes do VMware ESX e do ESXi. As principais interfaces de gerenciamento utilizadas pelos administradores do VMware ESX e do ESXi são:

- **VMware vSphere Client.** Permite gerenciar os hosts do VMware ESX ou do ESXi, as máquinas virtuais e, opcionalmente, o VMware vCenter Server com a interface comum de usuário do VMware vSphere Client. O vSphere Client está disponível para download gratuito e pode ser utilizado em um host do VMware ESX ou do ESXi para o gerenciamento de um único host ou utilizado no VMware vCenter Server para o gerenciamento de vários hosts.
- **VMware vCenter Server.** Possibilita gerenciar os hosts do VMware ESX e do ESXi e suas máquinas virtuais. Para gerenciar um host do ESX ou do ESXi com o VMware vCenter Server, é necessária uma licença do VMware vCenter Agent, incluída em todas as edições do VMware vSphere para esse host. O VMware vSphere possui vários outros recursos de gerenciamento que melhoram a continuidade dos negócios e maximizam a eficiência operacional, como migração em tempo real, balanceamento automático de carga, proteção contra falhas de hardware e recursos de backup e restauração de máquinas virtuais.
- **VMware vSphere Command-Line Interface 4.0 (vCLI).** Permite gerenciar o VMware ESX e o ESXi por meio de um ambiente de execução remota. A versão mais recente do vCLI possui uma série de novos comandos e é suportada tanto pelo VMware ESX 4.0 como pelo VMware ESXi 4.0.
- **VMware vSphere Power Command-Line Interface 4.0 (PowerCLI).** Permite gerenciar e configurar milhares de máquinas virtuais com esta interface fácil de usar, baseada na tecnologia Microsoft PowerShell. O PowerCLI permite que os administradores de TI gerenciem o VMware ESX ou o ESXi por meio de uma interface de scripts que gerencia as mesmas tarefas realizadas com o VMware vSphere Client.
- **VMware vSphere Management Assistant.** O VMware vSphere Management Assistant é uma máquina virtual que inclui a interface de linha de comando do VMware vSphere e outros softwares que os desenvolvedores e administradores podem usar para executar agentes e scripts, a fim de gerenciar os hosts do VMware ESX e do ESXi.
- **Gerenciamento de hardware sem agente com o CIM.** O CIM (*Common Information Model*, Modelo comum de informações) fornece um protocolo para monitoramento da integridade e do status do hardware, por meio de ferramentas de terceiros compatíveis com o CIM ou com o VMware vCenter Server.

8.7.7. Desempenho e Escalabilidade

O VMware ESX e o VMware ESXi possuem excelente desempenho e escalabilidade, segundo a VMware, permitindo que até mesmo os aplicativos de produção que mais utilizam recursos sejam virtualizados. Os principais recursos são:

- **Otimizações de desempenho para cargas de trabalho virtualizadas.** O VMware ESX e o ESXi 4.0 passaram por otimizações de desempenho para aplicativos específicos essenciais aos negócios, como Oracle Database, Microsoft SQL Server e Microsoft Exchange.
- **Desempenho aprimorado para armazenamento iSCSI.** Permite utilizar uma combinação entre os novos drivers SCSI otimizados para VIRTUALIZAÇÃO no SO convidado e as otimizações de pilha de armazenamento no nível do VMkernel, para aumentar significativamente o desempenho dos aplicativos de I/O intenso, como bancos de dados e aplicativos de mensagens.
- **Suporte a hardware de servidor potente.** Aproveita os sistemas de hardware com até 64 núcleos físicos de CPU, 256 CPUs virtuais, 1 TB de RAM e até centenas de máquinas virtuais em um único host, para facilitar a consolidação em larga escala e os projetos de recuperação de desastres.
- **Suporte a máquinas virtuais maiores.** Permite configurar máquinas virtuais com até 255 GB de RAM.
- **Suporte a SMP virtual de oito vias.** O VMware *Virtual Symmetric Multi-Processing* (SMP) melhora o desempenho das máquinas virtuais ao permitir que uma única máquina virtual utilize, de maneira simultânea, até oito processadores físicos. O VMware Virtual SMP permite a virtualização dos aplicativos corporativos que mais utilizam a CPU, como bancos de dados, ERP e CRM.
- **Mapeamento bruto de dispositivos.** Como opção, permite mapear LUNs de SAN diretamente para uma máquina virtual, a fim de ativar o cluster de aplicativos e a tecnologia de snapshot com base em array aproveitando, ao mesmo tempo, as vantagens da facilidade de gerenciamento do VMware vStorage VMFS.
- **Suporte para virtualização de hardware.** O VMware ESX e o ESXi fornecem suporte para as tecnologias de virtualização auxiliada por hardware de próxima geração, como o *Rapid Virtualization Indexing* da AMD ou o *Extended Page Tables* da Intel.
- **Suporte para páginas de memória amplas.** O VMware ESX e o ESXi são os únicos HYPERVISORS que suportam páginas de memória amplas, para aprimorar a eficiência do acesso à memória dos sistemas operacionais convidados.
- **Otimizações do desempenho de rede.** O VMware ESX e o ESXi suportam uma variedade de tecnologias de transferência de desempenho, como TSO (*TCP Segmentation Offloading*, segmentação TCP), VLAN e transferência de soma de verificação, e jumbo frames para reduzir a sobrecarga da CPU associada ao processamento de I/O da rede. Além disso, os recursos otimizados de desempenho de I/O da VIRTUALIZAÇÃO, como o NetQueue, são suportados, o que melhora

significativamente o desempenho em ambientes virtualizados Ethernet de 10 Gigabit.

- **Suporte para novos dispositivos e protocolos de alto desempenho.** O VMware ESX e o ESXi suportam as placas de rede Ethernet de 10 Gbps, os arrays de armazenamento e a tecnologia Infiniband para aprimorar o desempenho da máquina virtual.
- **Suporte para PARAVIRTUALIZAÇÃO.** O VMware ESX e o ESXi suportam sistemas operacionais convidados Linux paravirtualizados para aprimorar o desempenho da máquina virtual.
- **VMDirectPath de I/O para máquinas virtuais.** Permite aumentar a eficiência da CPU para aplicativos que exigem acesso frequente aos dispositivos de I/O, ao permitir que máquinas virtuais selecionadas acessem diretamente os dispositivos de hardware subjacentes. Outros recursos de virtualização, como o VMware vMotion, a independência de hardware e o compartilhamento dos dispositivos físicos de I/O, não estarão disponíveis para as máquinas virtuais que utilizam esse recurso.

8.7.8. Disponibilidade

O VMware ESX e o VMware ESXi oferecem alta disponibilidade para as máquinas virtuais no DATACENTER. Os principais recursos vinculados ao aumento da disponibilidade são:

- **Recurso de multipath incorporado para acesso ao armazenamento.** Permite garantir a disponibilidade do armazenamento compartilhado utilizando o recurso de multipath da SAN Fibre Channel ou SAN iSCSI.
- **Combinação de NICs.** Permite atribuir a cada máquina virtual de rede *failover* incorporado de NIC e balanceamento de carga, aumentando a disponibilidade e a tolerância a falhas do hardware. As políticas de combinação de NICs permitem que os usuários configurem vários adaptadores ativos e em standby.
- **Suporte para o Microsoft Clustering Services.** Permite implementar cluster de máquinas virtuais que executam o sistema operacional Microsoft Windows nos hosts físicos.

8.7.9. Interoperabilidade

O VMware ESX e o VMware ESXi são testados e certificados em toda a pilha de TI, que inclui servidores, sistemas de armazenamento, sistemas operacionais e aplicativos, permitindo a padronização em toda a empresa. A interoperabilidade é garantida em termos de:

- **Hardware de servidor.** O VMware ESX e o ESXi foram certificados para servidores instalados em RACK, torre e blade de líderes do setor como a Dell, HP, IBM, NEC, SUN e Unisys.

- **Hardware de armazenamento**. O VMware ESX e o ESXi foram certificados para uma ampla variedade de sistemas de armazenamento da Dell, EMC, Fujitsu, HP, Hitachi Data Systems, IBM, NEC, Network Appliance, Sun e 3PAR. Drivers SATA internos, DAS, NAS e SAN Fibre Channel e SAN iSCSI são suportados.
- **Sistemas operacionais**. O VMware ESX e o VMware ESXi suportam a maior variedade de sistemas operacionais não modificados, incluindo Windows, Linux, Solaris, Novell NetWare e outros.
- **Aplicativos de software.** Permitem executar qualquer aplicativo de software nas máquinas virtuais da VMware sem a necessidade de modificá-lo.
- **Formatos de máquinas virtuais**. O VMware ESX e o ESXi podem executar máquinas virtuais criadas em formatos não-VMware. Ao utilizar o VMware vCenter Converter gratuito, os usuários podem converter e executar máquinas virtuais do Microsoft Virtual Server e Virtual PC e do Symantec LiveState Recovery nos hosts do VMware ESX e do ESXi.

8.7.10. Segurança

Os recursos de segurança do VMware ESX e do ESXi protegem os dados armazenados no ambiente virtual. Os principais recursos são descritos aqui:

- **VMware VMsafe**. O VMware VMsafe é uma nova tecnologia de segurança que ajuda a proteger cargas de trabalho virtualizadas de uma maneira que antes não era possível com as máquinas físicas. O VMsafe fornece um conjunto de APIs de segurança que permite aos produtos de segurança de terceiros obter a mesma visibilidade do VMware ESX ou do ESXi durante a operação de uma máquina virtual para identificar e eliminar malware. Essa proteção avançada é obtida por meio da visibilidade granular dos recursos de hardware da máquina virtual, como memória, CPU e disco e I/O.
- **Proteção com o VMkernel**. O VMware ESX e o ESXi estão protegidos dos ataques e das explorações mais comuns, pois garantem a integridade do VMkernel, um componente essencial dos HYPERVISORS. As técnicas de integridade de disco no ESX e no ESXi protegem a inicialização do HYPERVISOR ao utilizar o TPM (*Trusted Platform Module*, módulo de plataforma confiável), um dispositivo de hardware integrado aos servidores. Os módulos do VMkernel, que são carregados no disco e na memória, recebem assinatura e validação digitais durante o carregamento, para garantir a autenticidade e integridade do código carregado dinamicamente e evitar tentativas de ataque de malware para modificar o VMkernel enquanto ele permanece no disco. O VMKernel também utiliza técnicas para manter a integridade da memória durante o carregamento combinadas com os recursos de microprocessador, para proteger a si mesmo de ataques comuns de estouro de buffer utilizados para explorar códigos em execução.
- **Criptografia.** Garante uma conexão segura aos hosts do VMware ESX e do ESXi com a criptografia SSL.

- **Permissão de autenticação para dispositivos iSCSI.** O VMware ESX e o ESXi protegem amplamente os dispositivos iSCSI contra intrusões ao exigir que o host ou o iniciador iSCSI seja autenticado pelo dispositivo iSCSI ou pelo iSCSI de destino sempre que o host tentar acessar os dados no LUN de destino.
- **Políticas de segurança de rede.** Permite garantir segurança para as máquinas virtuais na camada Ethernet. Impede a prática de *sniffing* em modo promíscuo do tráfego de rede, as alterações de endereços MAC e a falsificação de transmissões de MAC de origem.

8.8. Sistema de Arquivos vStorage VMFS

8.8.1. Introdução

O vStorage Virtual Machine File System (VMFS) é um sistema de arquivo do tipo cluster (CSF) para máquinas virtuais (VMs) que faz parte do VMware vSphere 4 e não está disponível como produto independente. O sistema VMFS permite que vários servidores ESX acessem simultaneamente o storage. Cada VM é encapsulada em um pequeno conjunto de arquivos e o VMFS é o sistema de arquivos.

O VMFS habilita a VIRTUALIZAÇÃO além dos contornos de um único sistema. O VMFS melhora o uso dos recursos quando permite que várias VMs compartilhem um pool de recursos no storage. O VMFS é o fundamento básico do VMware Thin Provisioning, VMware VMotion, VMware Distributed Resource Scheduler (VMware DRS), VMware High Availability (VMware HA) e VMware Storage VMotion (storage VMotion).

O VMFS permite ofertar serviços distribuídos com base na VIRTUALIZAÇÃO, incluindo:

- **Otimização de recursos distribuídos.** O VMFS permite que várias instâncias do VMware ESX acessem o armazenamento da mesma máquina virtual e, como resultado, as máquinas virtuais podem ser migradas de forma dinâmica e automática entre instâncias do VMware ESX, viabilizando:
  - A alocação dinâmica de recursos em *pools* de recursos.
  - A migração em tempo real das máquinas virtuais em execução entre servidores distintos.
- **Alta disponibilidade.** O VMFS gerencia os bloqueios em disco e reservas de SCSI, permitindo:
  - Fazer o cluster de máquinas virtuais com o Microsoft Clustering Services.
  - Fazer a reinicialização automática das máquinas virtuais em servidores físicos distintos.
- **Backup eficiente fora do host.** O VMFS permite que um servidor proxy faça backup do snapshot de uma máquina virtual, enquanto a máquina virtual realiza leituras/gravações em seu sistema de armazenamento.

A **Figura 8-9** ilustra o papel do VMFS.

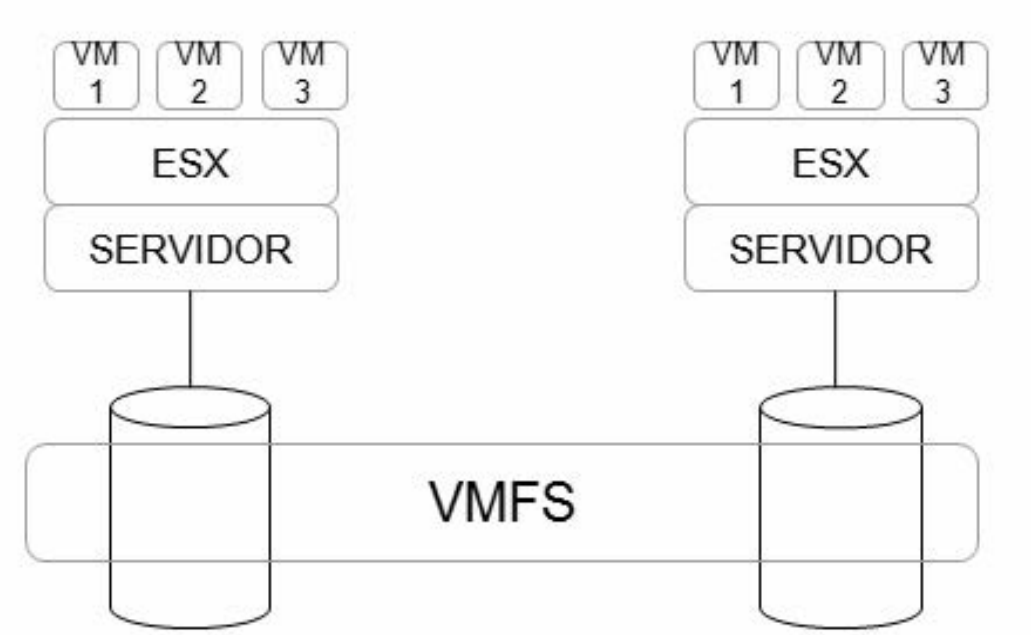


Figura 8-9 VMware vStorage VMFS

8.8.2. Funcionamento

O VMFS foi otimizado, testado e certificado para uma grande variedade de equipamentos Fibre Channel e SAN iSCSI. O VMFS armazena com eficiência todo o estado da máquina virtual em um local central e pode ser criado com antecedência, permitindo o provisionamento instantâneo de máquinas virtuais sem depender de um administrador de armazenamento.

Os sistemas de arquivos convencionais permitem que somente um servidor tenha acesso de leitura/gravação ao mesmo arquivo por vez. O VMFS é um sistema de arquivos que aproveita o armazenamento compartilhado para permitir a várias instâncias do VMware ESX acesso simultâneo de leitura/gravação ao mesmo sistema de armazenamento. O VMFS fornece bloqueio em disco para garantir que uma máquina virtual não seja inicializada por várias instalações do VMware ESX ao mesmo tempo. No caso de falha em um servidor, o bloqueio em disco de cada máquina virtual é liberado, permitindo que a máquina virtual seja reinicializada em outros servidores físicos.

8.8.3. Recursos

- **Sistema de arquivos de cluster.** Permite criar a base para os serviços de infraestrutura distribuída baseados na VIRTUALIZAÇÃO, armazenando os arquivos das máquinas virtuais em um sistema de armazenamento compartilhado, como Fibre Channel e SAN iSCSI.

- **Sistema de arquivos com dados compartilhados.** Permite que várias instalações do VMware ESX leiam e gravem simultaneamente no mesmo local de armazenamento.
- **Inserção ou exclusão de nós on-line.** Permite adicionar ou excluir um host do VMware ESX de um volume do VMFS sem pausar ou interromper o processamento de outros hosts do ESX no mesmo volume.
- **Bloqueio de arquivos em disco.** Evita que as máquinas virtuais sejam inicializadas por vários servidores ao mesmo tempo.
- **Otimizado para I/O de máquina virtual.** Permite armazenar e acessar o estado completo da máquina virtual com eficiência a partir de um local centralizado, com desempenho de disco virtual semelhante ao do SCSI nativo.
- **Dimensionamento adaptável de blocos.** Utiliza grandes blocos, aproveitando a I/O de disco virtual. Usa o distribuidor de sub-blocos para arquivos e diretórios pequenos.
- **Aumento dinâmico do tamanho dos volumes do VMFS.** Cria novas máquinas virtuais sem depender de um administrador de armazenamento. O dimensionamento adaptável de blocos e o tratamento de arquivos crescentes permitem aumentar um volume do VMFS de forma imediata.
- **Maior número de hosts do ESX por volume do VMFS.** Permite conectar até 32 hosts do ESX a um único volume do VMFS.
- **Limites ampliados de tamanho de bloco e arquivo.** Executa, nas máquinas virtuais, até mesmo os aplicativos de produção que mais consomem recursos, como bancos de dados, ERP e CRM.
- **Cache.** O cache de volume, dispositivo, objeto e buffer permite obter os benefícios do melhor desempenho do VMFS.
- **Certificação.** Permite utilizar o VMFS com uma grande variedade de equipamentos Fibre Channel e SAN iSCSI. O VMFS foi otimizado, rigorosamente testado e certificado para esses sistemas de armazenamento.
- **Discos virtuais compatíveis com SCSI.** Usa os arquivos de discos virtuais que aparecem para a máquina virtual como um dispositivo SCSI montado. Os discos virtuais ocultam todos os erros intermitentes de SAN do sistema operacional, permitindo que até mesmo sistemas operacionais não certificados para SAN sejam executados dentro de uma máquina virtual.
- **Detecção e gerenciamento de LUNs.** Simplifica o gerenciamento do armazenamento com a detecção automática de LUNs no sistema de armazenamento compartilhado e o mapeamento desses LUNs para um volume do VMFS.
- **Diretórios de arquivos.** Facilita a administração de máquinas virtuais com os diretórios de arquivos. Todos os arquivos de uma máquina virtual são armazenados em um diretório à parte.
- **Passagem direta dos dados da máquina virtual.** Garante o comportamento correto dos aplicativos e a integridade dos dados de aplicativos executados em má-

quinas virtuais. O VMFS preserva a semântica interna do sistema de arquivos do sistema operacional executado na máquina virtual.

- **Namespace hierárquico unificado.** Gerencia todos os discos físicos, volumes lógicos e volumes do VMFS disponíveis com um namespace consistente que elimine possíveis conflitos.
- **Snapshots de máquinas virtuais.** Aumentam a disponibilidade dos aplicativos e reduzem as janelas de backup usando snapshots das máquinas virtuais. Permitem criar uma cópia point-in-time dos dados da máquina virtual, que pode ser usada para operações de teste, backup e recuperação.
- **Adição dinâmica de disco virtual com o sistema em funcionamento.** Adiciona um disco virtual a uma máquina virtual em execução para aumentar os recursos disponíveis ou para backup.
- **Registro em log distribuído.** Permite recuperar máquinas virtuais de modo mais rápido e confiável no caso de falhas de servidor.

8.8.4. RDM

O VMware vSphere também permite a utilização de um sistema de arquivos chamado RDM (*Raw Device Mapping*). O RDM permite entregar uma LUN para a VM. Esta LUN vai ser formatada normalmente. Na verdade o RDM visa melhorar o desempenho de I/O, mas introduz uma maior dificuldade no gerenciamento. Para otimizar a performance de disco considere utilizar o RDM. O RDM permite que uma VM gerencie um disco bruto como um disco virtual VMFS.

8.9. vMotion

8.9.1. Introdução

A VMware desenvolveu uma série de funcionalidades que dependem da tecnologia vMotion.

O vMotion é um recurso básico para o funcionamento da alta disponibilidade (HA) e do balanceamento de carga (DRS) das soluções de virtualização da VMware. Essencialmente, o VMware VMotion permite a migração de máquinas virtuais em tempo real (migração viva).

O vMotion exige a utilização de um storage compartilhado por vários servidores em rede, onde a máquina virtual é encapsulada por um conjunto de arquivos armazenados no storage. A migração realizada com o vMotion permite, por exemplo, que máquinas virtuais se movam de servidores muito carregados para servidores ociosos.

8.9.2. Utilização

O VMotion permite:

- Otimizar e alocar automaticamente pools de recursos, possibilitando a utilização de hardware, flexibilidade e disponibilidade máximas.

- Realizar manutenção de hardware sem tempo de inatividade programado.
- Migrar máquinas virtuais de forma proativa, retirando-as de servidores com falhas ou baixo desempenho.

8.9.3. Funcionamento

A **Figura 8-10** ilustra o funcionamento do VMotion. A migração em tempo real de uma máquina virtual de um servidor físico para outro com o VMotion é possibilitada por três tecnologias.

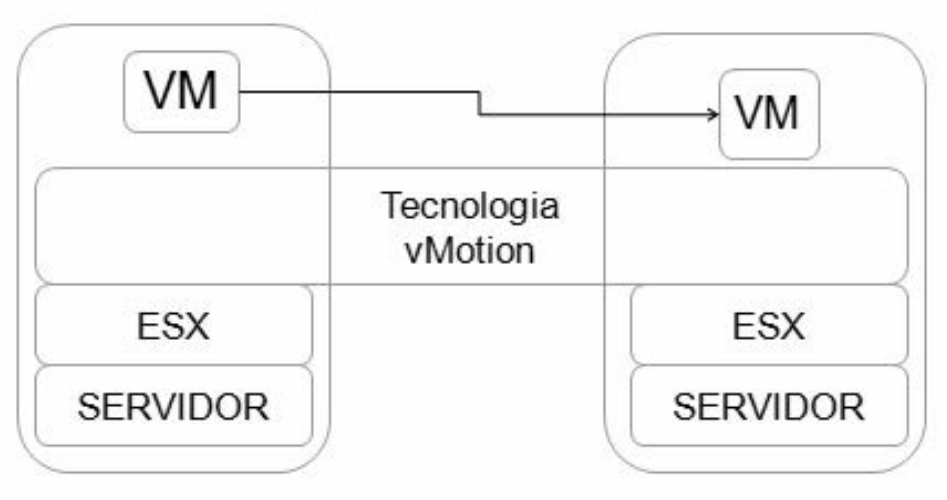


Figura 8-10 VMware vMotion

- O estado de uma máquina virtual é encapsulado por um conjunto de arquivos armazenados em um storage compartilhado em SAN Fibre Channel, SAN iSCSI ou NAS. O VMFS permite que várias instalações do ESX Server acessem simultaneamente os mesmos arquivos em uma máquina virtual.
- A memória ativa e o estado preciso de execução da máquina virtual são transferidos rapidamente por uma rede de alta velocidade, permitindo que a máquina virtual transfira instantaneamente sua execução do ESX Server de origem para o ESX Server de destino. O VMotion torna o período de transferência imperceptível para os usuários, mantendo o controle das transações contínuas de memória em um bitmap. Quando toda a memória e o estado do sistema tiverem sido copiados para o ESX Server de destino, o VMotion suspende a máquina virtual de origem, copia o bitmap para o ESX Server de destino e reinicia a máquina virtual no ESX Server de destino. Todo esse processo leva menos de dois segundos em uma rede Gigabit Ethernet, por exemplo.
- As redes em uso pela máquina virtual também são virtualizadas pelo ESX Server subjacente, garantindo que, mesmo após a migração, a identidade e as conexões de rede da máquina virtual sejam preservadas. O VMotion gerencia o endereço MAC virtual como parte do processo. Quando a máquina de destino é ativada, o VMotion notifica o roteador da rede, para garantir o reconhecimento da nova

localização física do endereço MAC virtual. Uma vez que a migração de uma máquina virtual com o VMotion preserva o estado preciso de execução, a identidade e as conexões ativas de rede, o resultado é tempo de inatividade zero e nenhuma interrupção para os usuários. Simplesmente fantástico.

8.9.4. Recursos

- **Confiabilidade.** Comprovado por clientes em ambientes de produção desde 2004, o VMotion é um padrão para os recursos de migração em tempo real.
- **Desempenho.** Realiza migrações em tempo real, com um tempo de inatividade imperceptível para os usuários finais. O uso otimizado da CPU e dos recursos de rede garante que as migrações em tempo real ocorram com rapidez e eficiência.
- **Interoperabilidade.** Permite migrar máquinas virtuais executando qualquer sistema operacional, em qualquer tipo de hardware e de storage compatível com o VMware ESX Server.
- **Suporte a SAN Fibre Channel.** Permite implementar a migração em tempo real de máquinas virtuais, utilizando uma grande variedade de sistemas de storage que utilizam SAN Fibre Channel.
- **Suporte a NAS e SAN iSCSI.** Permite implementar a migração em tempo real de máquinas virtuais com storage compartilhado de menor custo como SAN iSCSI e fácil gerenciamento.
- **Configurações personalizáveis de compatibilidade de CPU.** Garante que as máquinas virtuais possam ser migradas entre diferentes versões de hardware. Permite que as máquinas virtuais tirem proveito das mais recentes inovações de CPU.
- **Capacidade de Gerenciamento.**
  - **Assistente de migração.** Identifica o melhor destino para uma máquina virtual, utilizando informações em tempo real, fornecidas pelo assistente de migração.
  - **Várias migrações simultâneas.** Realiza várias migrações simultâneas para otimizar continuamente o posicionamento das máquinas virtuais em todo o ambiente de TI.
  - **Níveis de prioridade.** Atribui uma prioridade a cada operação de migração em tempo real, garantindo que as máquinas virtuais mais importantes tenham sempre acesso aos recursos de que precisam.
  - **Tarefas programadas de migração.** Automatizam as migrações para que ocorram em horários predefinidos e sem a presença de um administrador.
  - **Trilha de auditoria de migração.** Mantém um registro detalhado das operações de migração, incluindo a data/hora e os administradores responsáveis por iniciá-las.

8.10. vCenter Server

8.10.1. Conceito

O VMware vCenter Server fornece uma plataforma escalável e extensível para o gerenciamento da VIRTUALIZAÇÃO, permitindo ter uma visibilidade abrangente da infraestrutura virtual. O VMware vCenter Server gerencia de maneira centralizada os ambientes do vSphere e simplifica as tarefas diárias, melhorando o controle administrativo do ambiente.

O VMware vCenter Server é o aplicativo que gerencia o ambiente VMware vSphere. O vCenter Server fornece gerenciamento unificado de todos os hosts e máquinas virtuais de um DATACENTER a partir de um único console. Além disso, permite que os administradores tenham mais controle, simplifiquem as tarefas diárias e reduzam a complexidade e os custos de gerenciamento de um ambiente de TI.

8.10.2. Utilização

O VMware vCenter Server fornece uma plataforma escalável e extensível de gerenciamento da virtualização. Com o VMware vCenter Server, as organizações de TI obtêm:

- **Gerenciamento centralizado.** O vCenter Server permite que as empresas de TI organizem, provisionem e configurem com rapidez todo o ambiente de TI por meio de uma interface única, reduzindo os custos operacionais. O monitoramento preciso e consistente do desempenho de todos os componentes críticos– incluindo CPU, memória, armazenamento e rede – fornece detalhes de que os administradores precisam.
- **Automação operacional.** A programação de tarefas e de alertas melhora a agilidade de resposta às necessidades dos negócios e prioriza as ações que exigem atenção mais urgente.
- **Controle seguro de acesso.** Sólidos mecanismos de permissão e integração com o Microsoft Active Directory garantem acesso autorizado ao ambiente e às suas máquinas virtuais. As responsabilidades podem ser delegadas a administradores de sistemas de vários níveis.
- **Disponibilidade e gerenciamento de recursos.** Com o VMware vCenter Server, os administradores podem configurar e gerenciar o VMware vMotion, o Storage vMotion, o VMware High Availability (HA) e o Fault Tolerance (FT).
- **Níveis mais altos de segurança.** Aplicação automática de conformidade com padrões de patch pelo uso do VMware vCenter Update Manager, o que permite que as organizações protejam a infraestrutura virtual contra vulnerabilidades.
- **Eficiência de energia automatizada.** As empresas podem minimizar o consumo de energia com o VMware Distributed Power Management (DPM).

8.10.3. Funcionamento

A plataforma do VMware vCenter Server inclui vários componentes que funcionam em conjunto para fornecer às empresas um hub de gerenciamento de virtualização escalável.

Os servidores de gerenciamento fornecem pontos centrais de gerenciamento para os hosts e as máquinas virtuais, e as informações de inventário e desempenho são armazenadas em um banco de dados. Um agente vCenter fornece a conectividade entre o host e o servidor de gerenciamento.

Os administradores podem acessar o vCenter Server a partir do vSphere Client em qualquer PC com Windows ou usar o portal de acesso à Web do vCenter para acesso remoto a partir de qualquer navegador.

As funções e permissões são replicadas por todos os servidores de gerenciamento, concedendo aos administradores a capacidade de gerenciar vários vCenter Servers a partir de um único console. Um mecanismo de pesquisa torna possível encontrar rapidamente as máquinas virtuais, os hosts ou quaisquer outros objetos do inventário localizados em qualquer lugar da empresa.

8.10.4. Características

- Controle centralizado e visibilidade de todos os níveis da infraestrutura virtual:
  - **Monitoramento em tempo real dos elementos virtuais dinâmicos**. O vCenter Server reconhece os elementos virtuais e componentes físicos relacionados, incluindo hardware de servidor, armazenamento compartilhado e rede, com disparadores de eventos e alarmes que facilitam o monitoramento do ambiente, o diagnóstico e a solução de problemas. Os administradores podem visualizar as relações entre os servidores físicos, as máquinas virtuais, as redes e o armazenamento com mapas de topologia dinâmicos que verificam a configuração correta do vSphere. O vCenter Server monitora o desempenho e a disponibilidade das máquinas virtuais e dos outros elementos virtuais, como *pools* de recursos com estatísticas e gráficos detalhados que podem ser visualizados em tempo real.
  - **Disparadores de alarmes personalizáveis.** O vCenter Server é capaz de gerar notificações e alertas automatizados e disparar fluxos de trabalho automatizados para corrigir e prevenir problemas.
  - **Navegação e pesquisa simplificadas do inventário.** Utiliza a função de pesquisa global para acessar o inventário de vários vCenter Servers, incluindo máquinas virtuais, hosts, datastores e redes em qualquer lugar do vCenter. A interface de usuário permite uma navegação mais fácil.
- Gerenciamento proativo do VMware vSphere:
  - **vCenter Update Manager – Provisionamento rápido e gerenciamento de patches simplificado.** Cria novas máquinas virtuais ou hosts em questão de minutos, usando um assistente ou um modelo para minimizar erros e tempo de inatividade com os padrões de configuração. Faz correções de máquinas virtuais

e hosts com facilidade, usando o vCenter Update Manager, e padroniza e valida a configuração do host utilizando o recurso de perfis do host. Faz alocação dinâmica de recursos para garantir os SLAs. O vCenter Server monitora continuamente a utilização dos pools de recursos com o VMware DRS, que migra de modo inteligente as máquinas virtuais entre os hosts sem causar tempo de inatividade ou afetar os SLAs. O resultado é um ambiente de TI altamente otimizado e eficiente, que se gerencia sozinho, com balanceamento de carga incorporado. Quando as máquinas virtuais em um pool de recursos não necessitam de capacidade excedente, o VMware Distributed Power Management (DPM) coloca os hosts em standby para reduzir o consumo de energia sem afetar os SLAs.

- **vCenter Orchestrator – Automação do fluxo de trabalho.** O vCenter Server contém o vCenter Orchestrator, um mecanismo de orquestração que simplifica o gerenciamento ao permitir que os administradores automatizem mais de 800 tarefas por meio de fluxos de trabalho imediatos ou montem fluxos de trabalho com o uso de uma interface simples de arrastar-e-soltar.
- **Disponibilidade do vCenter Server.** Um console de gerenciamento de serviços exibe a integridade dos componentes do vCenter, permitindo que os administradores identifiquem e corrijam rapidamente problemas na infraestrutura de gerenciamento. O VMware vCenter Server Heartbeat (com licença separada) fornece disponibilidade adicional, reconhecendo todos os componentes do vCenter Server, realizando o failover do servidor de gerenciamento e do banco de dados na LAN ou na WAN para um servidor em standby.

- Escalabilidade e extensibilidade:
  - **Gerenciamento em larga escala.** O vCenter Server foi desenvolvido para lidar com os grandes ambientes de TI. Uma única instância do vCenter Server 4.0 gerencia até 300 hosts e 3.000 máquinas virtuais e, com o Linked Mode, é possível gerenciar até 1.000 hosts e 10.000 máquinas virtuais em 10 instâncias do vCenter Server, a partir de um único console.
  - **Arquitetura aberta.** As APIs do vCenter e uma extensão .NET permitem a integração entre o vCenter Server e outras ferramentas e suporta os plug-ins personalizados para o vSphere Client.

8.10.5. vCenter Server Heartbeat

O VMware vCenter se tornou um componente crítico das instalações de vSphere. O aumento da sua disponibilidade passou a ser uma consideração importante. A VMware com o vCenter Server Heartbeat resolveu esta questão.

Utilização

O VMware vCenter Server Heartbeat fornece alta disponibilidade e recuperação de desastres para o VMware vCenter Server e todos os seus componentes com failover na LAN ou na WAN, incluindo o banco de dados e o servidor de licença. O software suporta failover P2V (*Physical-to-Virtual*, físico para-virtual), P2P (*Physical-to-Physical*, físico-

-para-físico) e V2V (*Virtual-to-Virtual*, virtual-para-virtual), garantindo o funcionamento contínuo da infraestrutura VMware, quando o VMware vCenter Server estiver sob risco de passar por tempo de inatividade planejado ou não planejado. O software monitora o VMware vCenter Server e oferece ampla proteção contra erros de aplicativos ou operadores, falha de sistema operacional ou hardware ou eventos externos.

O VMware vCenter Server Heartbeat é fácil de instalar e configurar. O software não possui dependências de configuração de hardware e detecta automaticamente os componentes padrão do VMware vCenter Server na instalação para monitoramento e proteção imediatos.

O VMware vCenter Server Heartbeat permite aos administradores fornecer uma infraestrutura de DATACENTER virtual disponível e resiliente. Com o VMware vCenter Server Heartbeat, os administradores podem:

- Proteger e monitorar a disponibilidade do VMware vCenter Server, minimizando o risco de tempo de inatividade para o vCenter Server e garantindo acesso a componentes essenciais, como o VMware VMotion e o VMware DRS, além de monitorar e analisar funções que dependem do VMware vCenter Server.
- Estender o monitoramento e a alta disponibilidade do serviço para os componentes ou plug-ins do VMware vCenter Server, incluindo o VMware License Server, o VMware vCenter Update Manager e até mesmo as ferramentas terceirizadas que servem como plug-ins do VMware vCenter Server.
- Automatizar o failover e o failback do VMware vCenter Server e de todos os componentes em uma LAN ou WAN, eliminando paralisações caras e complexas com failover e failback sem interrupções.
- Proteger e recuperar a instância de banco de dados do VMware vCenter Server, mesmo que ele esteja instalado em um servidor diferente do que hospeda o VMware vCenter Server.

Funcionamento

O VMware vCenter Server Heartbeat protege o Vmware vCenter Server de três maneiras:

- **Replicação**: o VMware vCenter Server Heartbeat replica continuamente as informações do VMware vCenter Server para um servidor em espera passivo, a fim de realizar uma recuperação rápida. Os locais dos dados, dados de licença, bancos de dados e das entradas de registro são clonados automaticamente. O VMware vCenter Server Heartbeat utiliza um canal de rede oculto para manter a conectividade e replicar para um servidor secundário passivo, garantindo que apenas o servidor principal seja visível aos usuários.
- **Monitoramento**: o VMware vCenter Server Heartbeat utiliza tecnologia de previsão para monitorar continuamente a integridade do VMware vCenter Server, do banco de dados e de outros componentes ou plug-ins. O software monitora componentes, como o VMware License Server, o serviço vpxd, o VMware vCenter

Update Manager e outras peças essenciais da plataforma de gerenciamento. O monitoramento também pode ser facilmente estendido para qualquer plug-in do VMware vCenter Server.

- **Recuperação automatizada:** quando houver uma ameaça à disponibilidade, ou simplesmente uma janela de manutenção planejada, os administradores poderão transferir o VMware vCenter Server e todos os seus componentes de um servidor principal para um servidor secundário com apenas o clique de um botão. Todo o processo de failover e failback também pode ser facilmente automatizado, baseado em critérios predefinidos.

A **Figura 8-11** ilustra o vCenter Server Heartbeat.

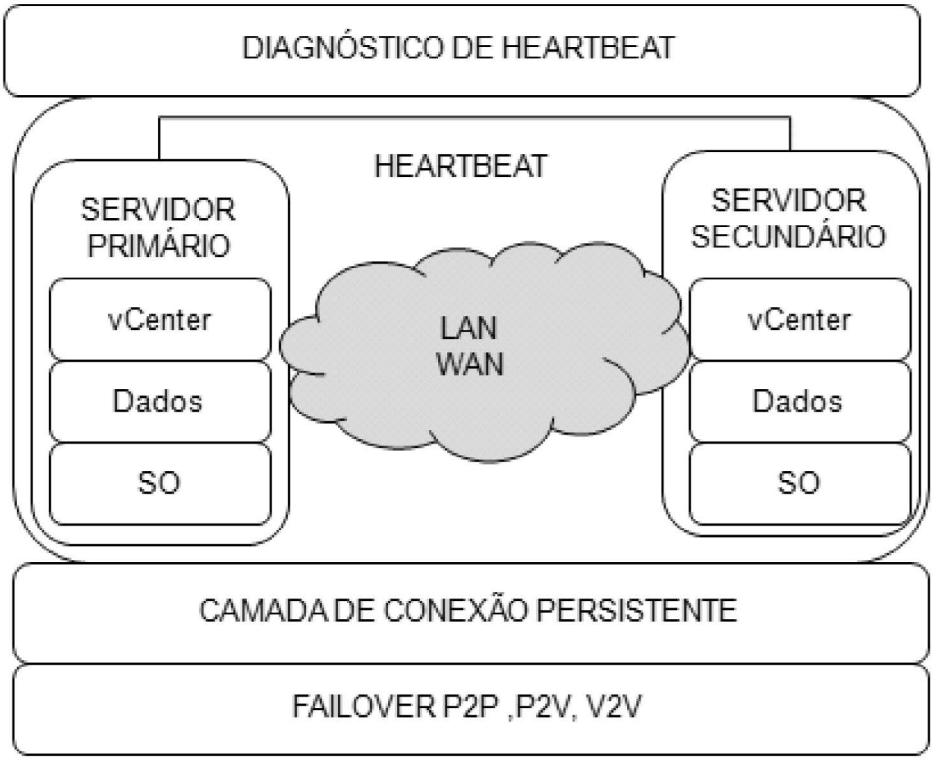


Figura 8-11 vCenter Server Heartbeat

8.10.6. vCenter Site Recovery Manager (SRM)

O VMware Site Recovery Manager (SRM) é uma solução de automação e gerenciamento de recuperação de desastres integrada ao VMware vSphere, ao VMware vCenter e ao software de replicação do storage.

O VMware SRM permite a obtenção dos níveis requeridos de RTO e RPO e pode também ser utilizado para efetuar a recuperação de dois ambientes com WORKLOADS ativas. O VMware SRM é integrado ao gerenciamento do VMware vCenter e deve operar integrado com o software utilizado para a replicação dos volumes no storage.

O VMware SRM permite:

- Gerenciar os planos de recuperação de desastres.
- Automatizar o failover e a recuperação.
- Simplifica e automatizar as cargas de trabalho envolvidas na recuperação de desastres.
- Permite fazer a configuração e testes de failover e failback.
- Possibilita fazer o gerenciamento central de planos de recuperação integrado ao VirtualCenter.
- Permite transformar os processos manuais de recuperação em planos de recuperação automatizados.
- Simplifica a integração com a replicação de armazenamento de terceiros.

A **Figura 8-12** ilustra o funcionamento do SRM.

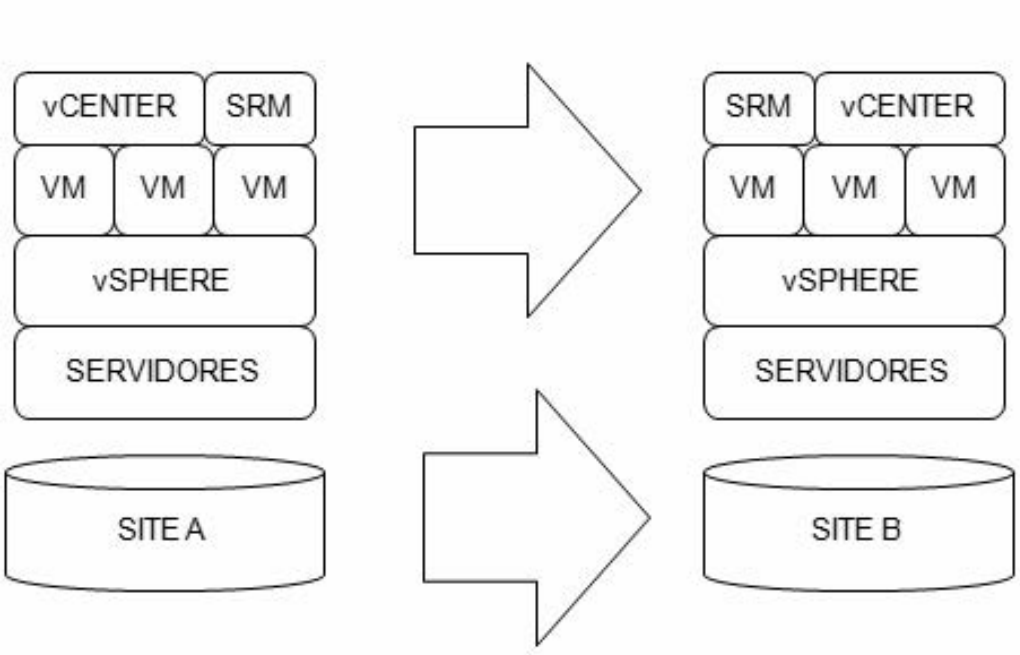


Figura 8-12 vCenter SRM

O VMware SRM pode ser utilizado também para recuperação de vários sites, conforme ilustra a **Figura 8-13**.

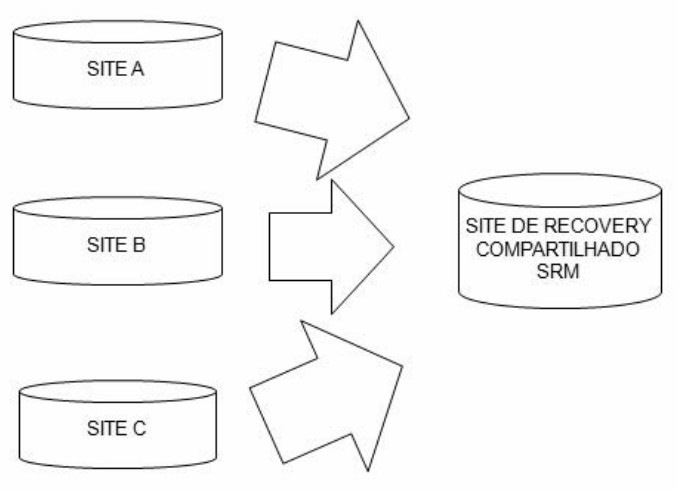


Figura 8-13 Opção de recuperação de site com o SRM

8.11. Funcionalidades Importantes do VMware vSphere

8.11.1. VMware High Availability (HA)

O HA do VMware permite obter alta disponibilidade para as máquinas virtuais, independentemente do hardware ou sistema operacional utilizado. Uma máquina virtual (VM) que roda em um servidor que venha a falhar é reinicializada automaticamente em outro servidor. Se um servidor de produção falha, as máquinas virtuais são automaticamente reinicializadas em outro servidor de produção. A **Figura 8-14** ilustra o uso do HA.

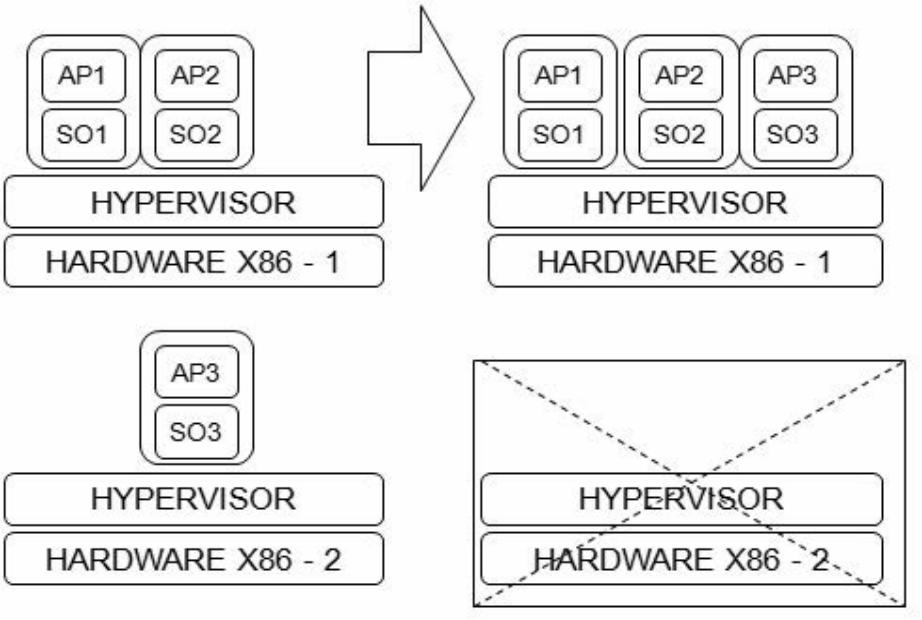


Figura 8-14 VMware HA

8.11.2. VMware Distributed Resource Scheduler (DRS)

O Distributed Resource Scheduler, do VMware, permite que as máquinas virtuais possam ser redistribuídas para outros servidores mediante algumas regras previamente estabelecidas que visam normalmente a melhoria da performance.

O DRS aloca e distribui dinamicamente no hardware a WORKLOAD das máquinas virtuais. O DRS monitora dinamicamente a WORKLOAD das máquinas virtuais que estão rodando e a utilização dos recursos em servidores físicos. O DRS verifica o uso dos recursos considerando as políticas predefinidas pelo administrador.

O *Performance Study DRS Performance and Best Practices* da VMware é uma excelente referência sobre a otimização da performance com o DRS.

A **Figura 8-15** ilustra o uso do DRS. É comum utilizar o recurso de DRS focado em otimização dos recursos de processamento e memória em conjunto com recursos de storage como o NQM do EMC Clariion que otimizam o I/O. Esta é uma opção interessante.

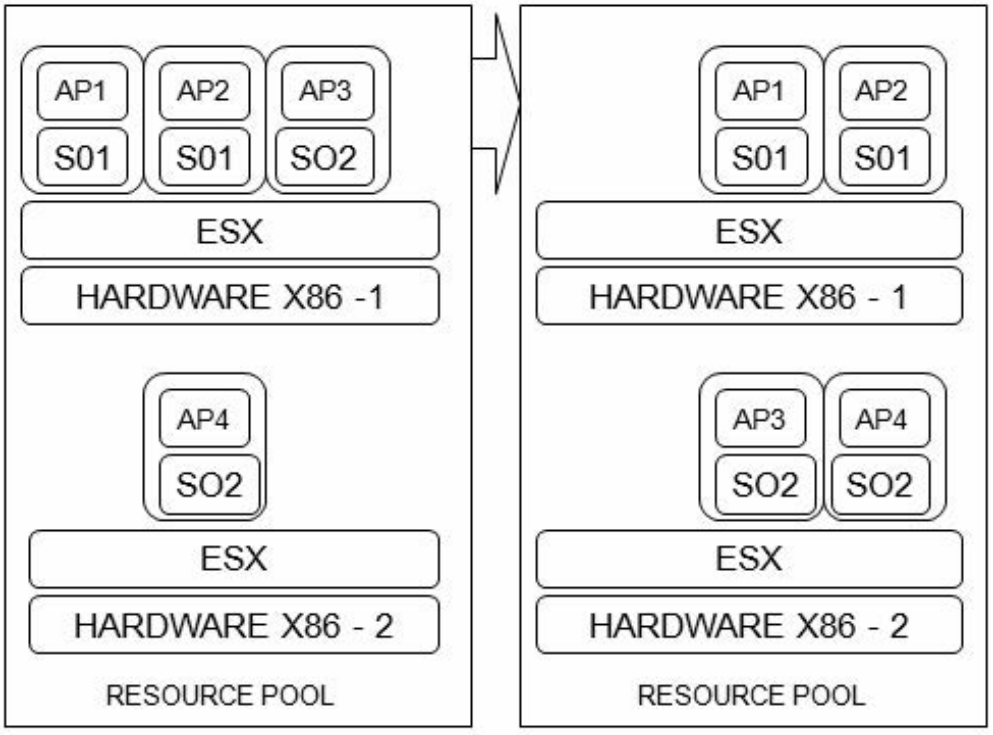


Figura 8-15 VMware DRS

8.11.3. VMware Distributed Power Management (DPM)

O consumo dos DATACENTERS nos Estados Unidos chegou a 1,5% do total da energia consumida, conforme descrito no capítulo 3. Os servidores consomem boa parte da energia consumida pela carga de TI. O VMware DRS usa a funcionalidade VMware Distributed Power Management (DPM) para reduzir o consumo de energia dos servi-

dores onde existe capacidade ociosa. Os servidores são trazidos de volta quando a demanda sobe novamente. O VMware DRS também reserva capacidade para o *failover* automático. Estas funcionalidades tornam a solução com o VMware muito mais flexível do que qualquer solução física.

8.11.4. VMware Fault Tolerance (FT)

A Consolidação de Servidores é uma realidade. A melhoria no uso de recursos e o menor consumo de energia são aspectos chaves resolvidos pela consolidação. A queda de um servidor físico em uma consolidação que não usa a VIRTUALIZAÇÃO pode ser menos problemática. Mas com a VIRTUALIZAÇÃO, os servidores físicos passam a ser críticos, principalmente devido ao uso de altas taxas de consolidação das soluções atuais.

O que acontece se o servidor físico que consolida diversas aplicações virtuais cair? O VMware HA consegue mover as aplicações de um servidor fisico para outro, mas no caso de haver um problema com uma máquina virtual específica não resolve o problema.

No mundo físico a tolerância a falhas de aplicações críticas é resolvida com hardware específico que normalmente é muito oneroso ou com a tecnologia de cluster via software, com o Microsoft Cluster Services (MSCS). Softwares de cluster normalmente demandam um servidor standby com uma configuração idêntica ao servidor ativo. O servidor standby deve ter uma cópia das aplicações e sistema operacional do sistema ativo. A aplicação ou as aplicações devem estar prontas para o cluster (*cluster aware*) para reduzir o nível de interrupção. A configuração do software de cluster pode ser complexa.

Em uma solução virtualizada utilizar uma técnica de cluster ou hardware específico pode não ser interessante, pois a complexidade seria trazida para o meio virtual.

O VMware Fault Tolerance (FT) constrói uma solução de alta disponibilidade integrada ao HYPERVISOR. O FT não requer configurações customizadas do sistema operacional convidado ou das aplicações. O FT estabelece e mantém uma cópia secundária da máquina virtual principal. A cópia secundária da máquina virtual reside em outro servidor físico e executa exatamente a mesma sequência de instruções da máquina virtual principal. Ambas as máquinas rodam em servidores separados, mas são gerenciadas como uma única entidade. A máquina secundária entra em operação de forma automática com a interrupção do serviço da máquina principal.

A **Figura 8-16** ilustra o funcionamento do FT.

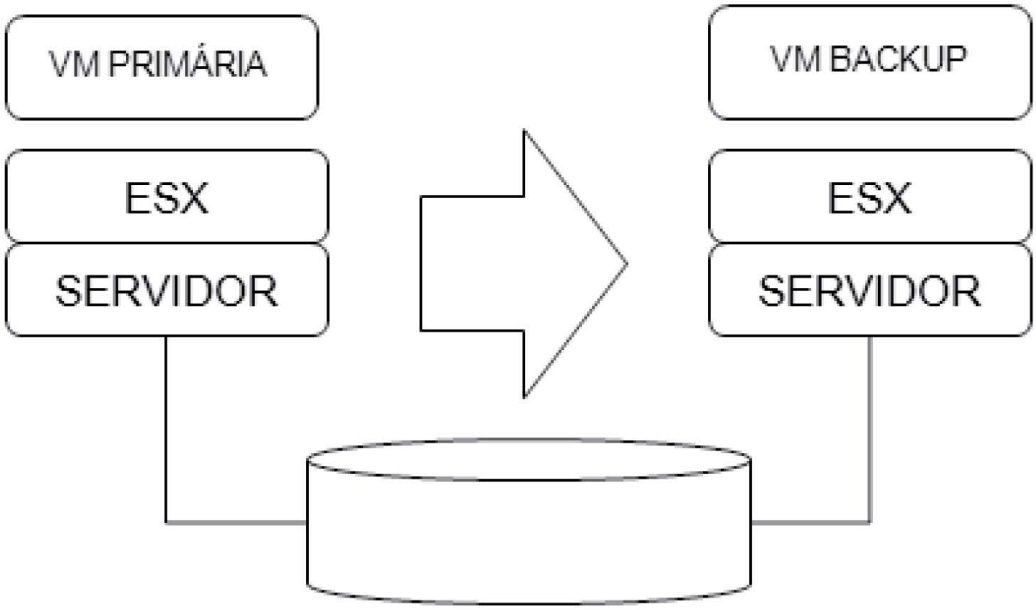


Figura 8-16 VMware Failover Transparente (FT)

O VMware FT preserva as transações sem qualquer perda garantindo a arquitetura de CPU, memória e I/O. As duas tecnologias chaves do FT são o vLockstep e o failover transparente.

A tecnologia vLockstep foi desenvolvida para garantir à arquitetura que os estados das máquinas primárias e secundárias sejam idênticos em qualquer ponto da execução das instruções. A tecnologia vLockstep garante que as instruções estão sendo executadas de forma exatamente igual nas máquinas virtuais principal e secundária. Ela requer extensões do processador físico desenvolvidas em conjunto por AMD e Intel e é integrada com o HYPERVISOR ESX. A tecnologia vLockstep enxerga uma única máquina virtual executando uma carga de trabalho. O HYPERVISOR, rodando nas duas máquinas físicas, estabelece um sistema de heartbeat e monitoramento mútuo quando inicia o vLockstep.

O ESX com o HA já é protegido contra falhas de componentes como falhas de HBAs SCSI ou interface de rede. Agora, se as duas HBAs do servidor principal falham ao mesmo tempo e as HBAs do servidor secundário continuam a funcionar, o FT providencia o failover.

O FT para ser habilitado em uma máquina virtual deve fazer parte de um cluster HA. O FT se baseia no serviço HA para criar um novo servidor como parte do failover transparente.

8.11.5. VMware Data Recovery (VDR)

Backup de máquina virtual é uma grande questão. O VMware Consolidated Backup (VCB) bem que tentou resolver a questão mas não conseguiu e foi retirado de linha. Agora é a vez do VMware Data Recovery (VDR).

O VDR permite realizar o backup full de máquinas virtuais e fazer a recuperação de máquinas virtuais de forma full ou incremental e a recuperação individual de aquivos ou diretórios. O VDR é baseado em snapshots e não utiliza agentes.

O VDR é constituído de três partes:

- A interface de usuário, que é um plug-in instalado dentro do vCenter.
- O VMware VDR, que é um appliance que responde pelo processo de backup e restore e utiliza a desduplicação.
- Um storage de destino.

A **Figura 8-17** ilustra o backup com o VDR em três etapas:

1. O backup é agendado via vCenter Server.
2. É criado o snapshot.
3. O dado é desduplicado e armazenado.

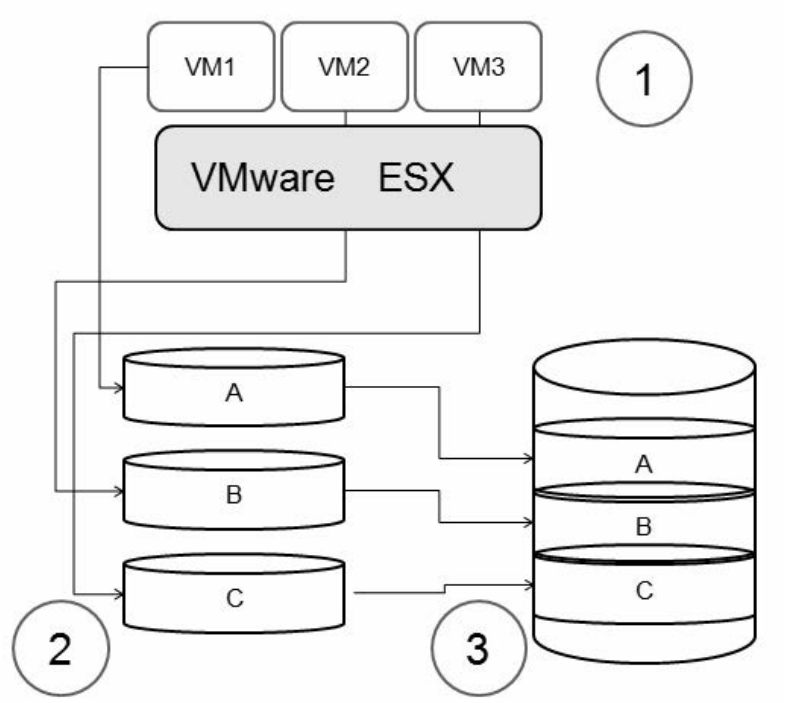


Figura 8-17 Backup com VDR

Características:

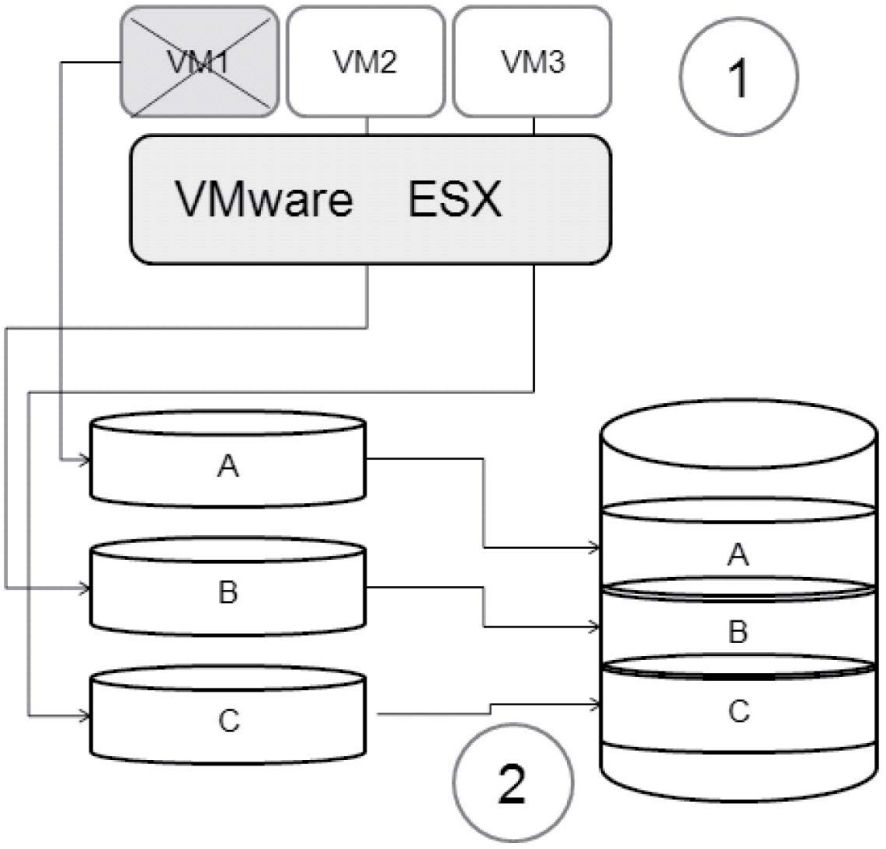
- O VDR é uma aplicação de backup com interface GUI integrado com o vCenter.
- O VDR é uma ferramenta de backup/restore para máquinas virtuais convidadas.
- O VDR é baseado em disco e permite backup full ou incremental rápido das VMs e de arquivos e diretórios para máquinas virtuais rodando o Windows como sistema operacional.

- O VDR usa naturalmente a desduplicação e é compatível com qualquer storage baseado em FC, iSCSI, NAS ou mesmo um storage local.
- Suporta o Volume Shadow Copy Services (VSS) para suportar o backup de aplicações Microsoft.

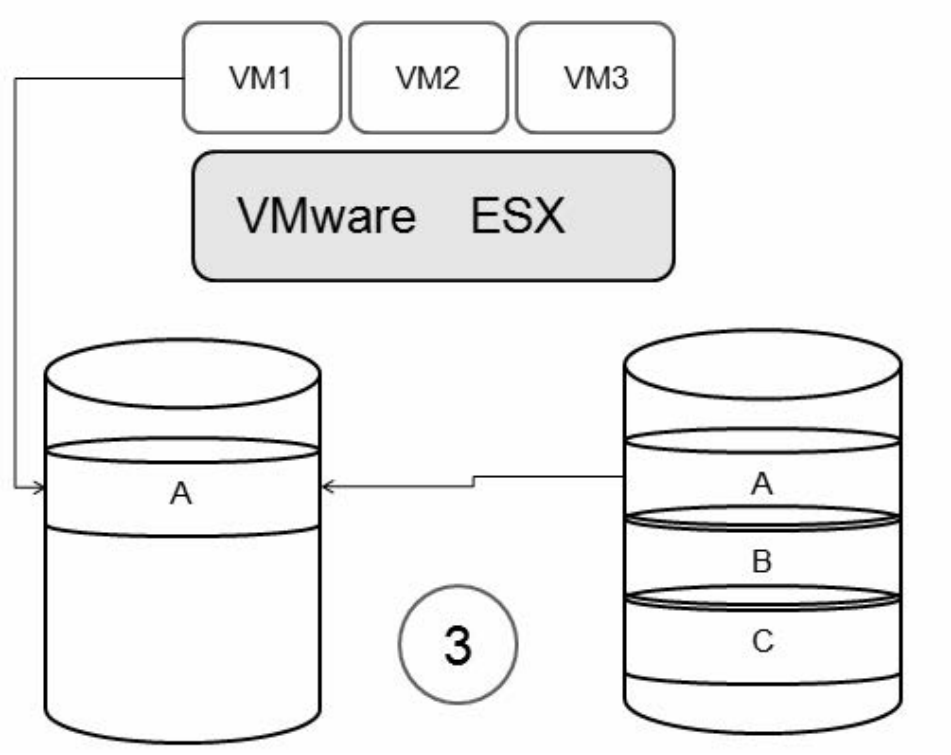
A grande limitação do VDR é que ele não envia o backup para a fita. E só faz backup de máquina virtual, e não de servidor físico. O VDR é um *appliance* no formato OVF. O número máximo de backup de VMs é 100. O VDR só pode ser utilizado com a versão ESX 4.0.

A **Figura 8-18 (a) e (b)** ilustra o restore com VDR.

1. VM cai.
2. Seleciona-se a VM ou os arquivos a recuperar.
3. Recupera-se a VM.



(a)



(b)

Figura 8-18 Recovery com o VMware VDR

O restore pode ser realizado também de um arquivo individual.

8.11.6. VMware Storage VMotion

O Storage VMotion do VMware realoca os arquivos de discos das máquinas virtuais entre unidades de storage diferentes com zero de downtime.

Requisitos:

O Storage VMotion requer a instalação do VMware vCenter Server e do vCenter agent.

O VMware Storage VMotion realoca os arquivos de discos das máquinas virtuais entre unidades de storage diferentes com zero de downtime.

O VMware Storage VMotion permite:

- A migração de dados de um RAID para outro que possa propiciar melhor desempenho.
- Que dados sejam movidos entre protocolos como FC ou iSCSI.

O Storage VMotion requer a instalação do VMware vCenter Server e do vCenter agent. A **Figura 8-19** ilustra o uso do Storage VMotion. Em T_1, as duas unidades de storage A e B funcionam normalmente. Em T_2, a unidade B apresentou um problema e a LUN B foi migrada pelo storage VMotion para a unidade de storage A.

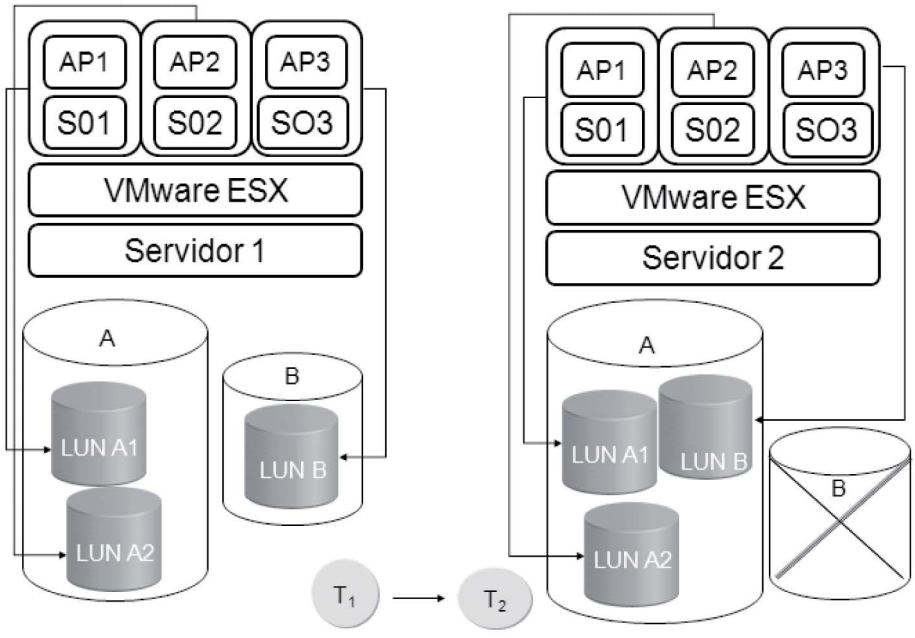


Figura 8-19 VMware Storage VMotion

8.11.7. VMware View

O VMware View 4.5 é uma solução de virtualização de *desktop* que proporciona uma opção de desktop virtual para usuários finais, tanto online como offline. Segundo a VMware, o View 4.5 é uma solução completa de desktop virtual que permite às empresas melhorar a segurança, reduzir custos operacionais e simplificar a administração da área de trabalho e estabelece uma arquitetura de computação moderna para o usuário final.

Benefícios

- Redução dos custos operacionais. Esta redução pode chegar a 50%, segundo o IDC, com o VMware View.
- Simplifica o gerenciamento de desktops virtuais, aplicações e dados de usuário de uma interface de gerenciamento centralizada, enquanto simplifica os processos de TI como o provisionamento e o gerenciamento da configuração.
- Nova experiência para usuários com a tecnologia PC over IP (PCoIP) com conexão LAN ou WAN, enquanto o VMware View Client, com o modo local, permite o trabalho offline.

- Estende a continuidade do negócio e a recuperação de desastres para o desktop.
- Padroniza a plataforma de VIRTUALIZAÇÃO.

A visão da VMware para a computação do usuário final é ajudar as empresas a migrar seus ambientes de computação de *desktop* para um modelo de aplicativos centrados no usuário.

Características e vantagens do VMware View 4.5:

- Acesso offline ao desktop.
- Migração de forma rápida para Windows 7.
- Recursos Administrativos expandidos.
- Distribuição de aplicativos integrados.
- Experiência superior para o usuário.
- Redução dos custos.
- Recursos de segurança avançados.
- Dispositivos *end-point* flexíveis.
- Proteção otimizada de antivírus.

Componentes do VMware View 4.5

- *VMware vSphere for Desktops*. Plataforma para rodar desktops virtuais e aplicações.
- *VMware View Manager*. Permite o gerenciamento centralizado dos desktops virtuais.
- *VMware ThinApp*. Aplicação sem agente que simplifica tarefas repetitivas e reduz a necessidade de storage para os desktops virtuais.
- *VMware Composer*. Permite criar imagens de desktops virtuais que compartilham discos virtuais com uma imagem única pai, enquanto segmenta os dados e as aplicações de usuário para que possam ser gerenciados de forma independente.
- *VMware vCenter for Desktops*. Versão do vCenter para gerenciamento do ambiente vSphere.
- *VMware View Client*. Permite o acesso dos PCs Windows aos desktops virtuais no DATACENTER.
- *VMware vShield Endpoint*. Solução centralizada offload de antivírus e antimalware.

Versões

O VMware View 4.5 é comercializado em duas versões que incluem:

- **VMware View Enterprise Edition**: inclui o VMware vSphere para Desktops, o VMware vCenter para Desktops e o VMware View Manager.

- **VMware View Premier Edition**: inclui o VMware vSphere para desktops, o VMware vCenter para Desktops, o VMware View Manager, o VMware View Composer, o VMware vShield Endpoint, o VMware ThinApp e o VMware View Client com Local Mode.

A **Figura 8-20** ilustra os componentes do VMware View.

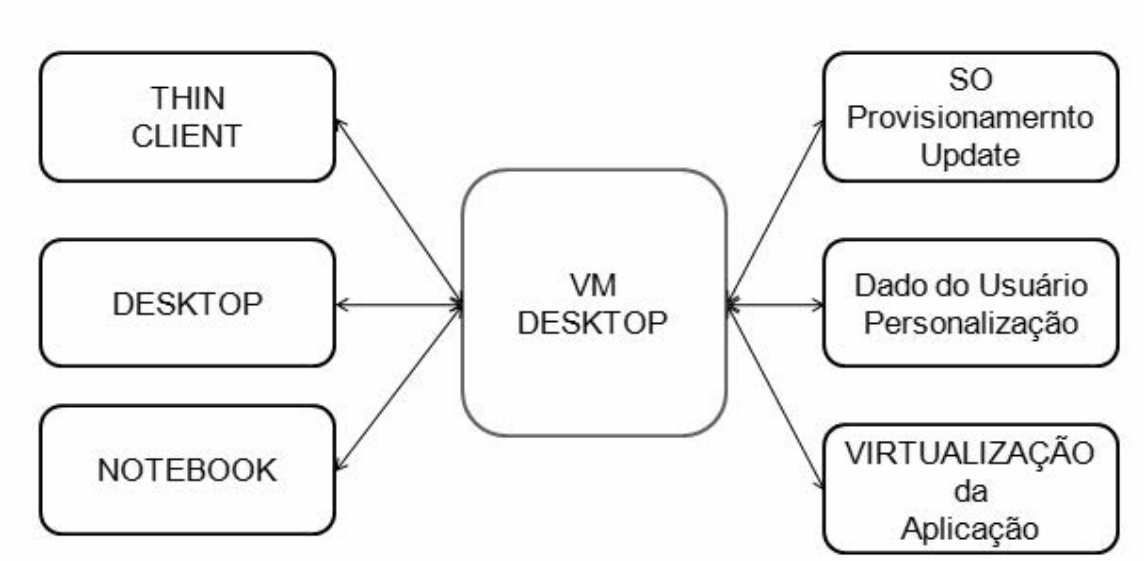


Figura 8-20 VMware View

8.12. Referências Bibliográficas

Haletky, Edward L. VMware ESX Server in the Enterprise: Planning and Securing Virtualization Servers. Prentice Hall PTR, 2008.

How VMware Virtualization right-sizes IT Infrastructure to Reduce Power Consumption, VMware, 2009.

iSCSI Design Considerations and Deployment Guide VMware Infrastructure 3, Technical Note,VMware, 2007.

Protecting Mission-Critical Workloads with VMware Fault Tolerance, VMware, 2009.

SAN System Design and Deployment Guide Second Edition, VMware July 2008.

The architecture of VMware ESXi, VMware, 2007.

VMware and VMotion and CPU Compatibility, Information Guide, VMware, 2007.

VMware Data Recovery, datasheet, VMware, 2010.

VMware ESX e Vmware ESXi: Os HYPERVISORS líderes de mercado com produção comprovada, VMware, 2009.

VMware ESX Server using EMC Clariion Storage Systems, EMC, 2007.

VMware vCenter Heartbeat: Garante disponibilidade para a plataforma de gerenciamento do VMware vCenter Server. Datasheet do produto, VMware, 2009.

VMware vCenter Server: Unifica e simplifica o gerenciamento da virtualização. Datasheet do produto, VMware, 2009.

VMware View 4.5, datasheet, VMware, 2010.

VMware Virtual Networking Concepts, Information Guide, VMware, 2007.

VMware VMFS vStorage Virtual Machine File System, Datasheet do produto,VMware, 2010.

VMware vSphere: The Best Platform for Cloud Infrastructure, VMware, 2009.

VMware vSphere 4: A melhor plataforma para a criação da infraestrutura em nuvem. Datasheet do Produto, VMware, 2009.

VMware vSphere, the First CLOUD Operating System, Provides an Evolutionary, Non-disruptive Path to CLOUD COMPUTING, VMware, 2009.

What is New in VMware vSphere 4: Virtual Networking, VMware, 2009.

What is New in VMware vSphere 4: Storage, VMware, 2009.

PARTE IV: PROJETO

9. Metodologia e Ferramentas de Projeto

9.1. Introdução

Um projeto é um esforço temporário empreendido para criar um produto, serviço ou resultado exclusivo. A implantação de um ambiente virtualizado pode ser considerada um projeto. Na verdade, pode-se assumir que existem dois projetos: um que define a ida da organização para a VIRTUALIZAÇÃO e escolhe o fornecedor, e outro que trata do projeto e implementação da solução. A complexidade envolvida sugere o uso de técnica de gestão de projetos.

A gestão de projetos é a aplicação de conhecimentos, habilidades, ferramentas e técnicas às atividades do projeto, a fim de atender aos seus requisitos. É realizada através de processos de gerenciamento de projetos sugeridos pelo PMBOK que são os de iniciação, planejamento, execução, controle e encerramento, conforme ilustra a **Figura 9-1**.

Os principais passos a serem dados em qualquer projeto são resumidos a seguir:

- **Iniciação**: abrir o projeto.
- **Planejamento**: desenvolver o plano de gerenciamento do projeto.
- **Execução**: orientar e gerenciar a execução do projeto.
- **Controle**: monitorar e controlar o trabalho do projeto.
- **Encerramento**: encerrar o projeto.

Deve-se em um projeto:

- Justificá-lo através de um business case ou atendendo a uma demanda do planejamento estratégico.
- Abri-lo formalmente através da carta de abertura.
- Identificar os requisitos, ou seja, as necessidades do projeto, gerando a documentação de requisitos.
- Estabelecer objetivos claros e alcançáveis e fazer a declaração do escopo.

- Balancear as demandas conflitantes de qualidade, tempo, escopo e custo.
- Adaptar as especificações, os planos e a abordagem às diferentes preocupações e expectativas das diversas partes interessadas.
- Concluir o projeto, fechar contratos e preencher a carta de encerramento explicitando as lições aprendidas.

Na **Figura 9-1** ilustram-se também as duas formas comuns de justificar um projeto: pela realização de um business case ou por uma demanda a ser atendida definida no plano estratégico.

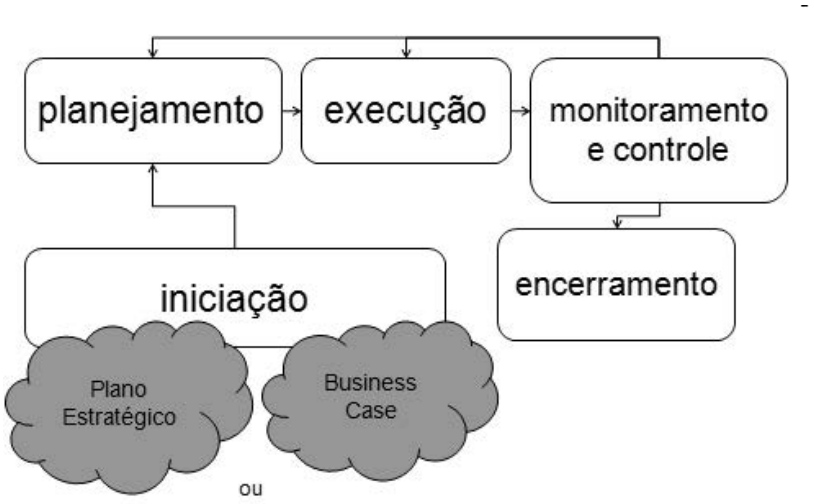


Figura 9-1 Processos de Gerenciamento de Projetos

Os principais documentos que devem ser gerados durante o ciclo de iniciação e planejamento do projeto de VIRTUALIZAÇÃO são ilustrados na **Figura 9-2**.

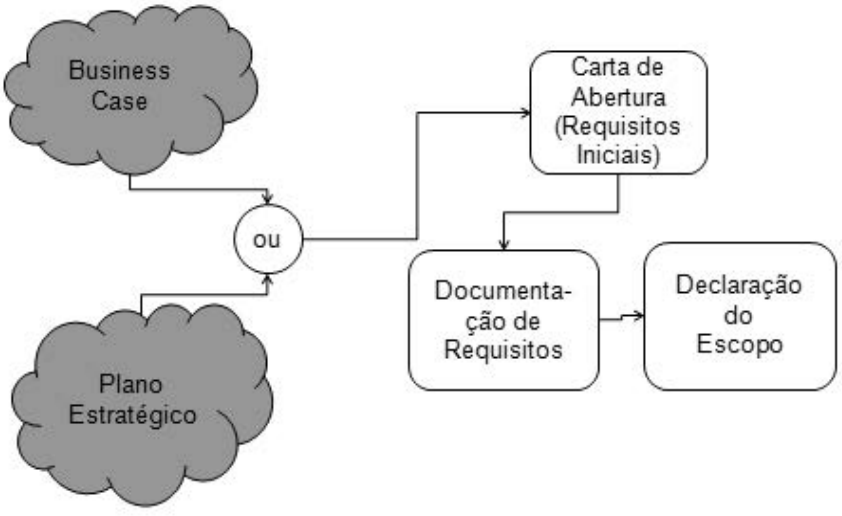


Figura 9-2 Documentos gerados durante o Ciclo de Gerenciamento de Projeto

9.2. Gestão de Projetos de VIRTUALIZAÇÃO

O projeto do DATACENTER pode ser dividido em duas grandes partes: o projeto da infraestrutura de TI, que tem como núcleo a VIRTUALIZAÇÃO, e o projeto das instalações, que depende do projeto da infraestrutura de TI. Na prática, precisa-se saber o tamanho da carga de TI para efeito do dimensionamento das instalações, energia necessária e estratégia de refrigeração.

A **Figura 9-3** ilustra as etapas do projeto do DATACENTER.

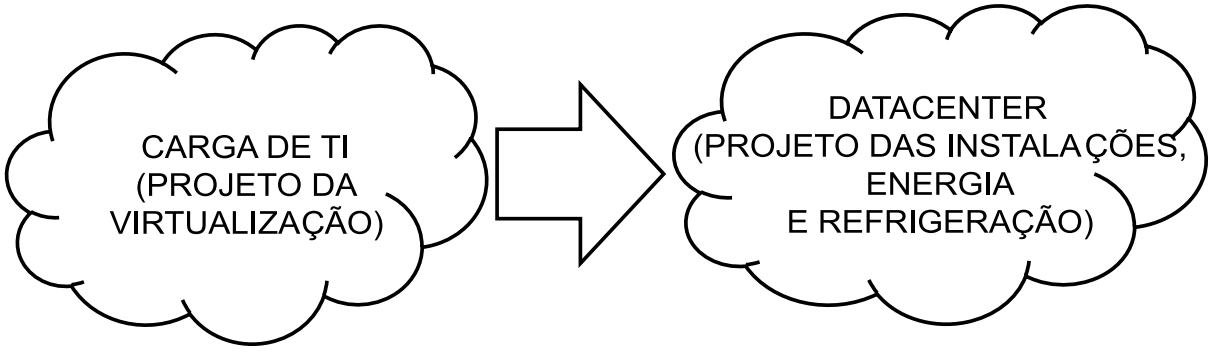


Figura 9-3 Projeto do DATACENTER

9.2.1. Projeto do DATACENTER

O projeto do DATACENTER inicia com a definição da energia necessária para cada RACK específico. Recomenda-se definir RACKS com consumos padrões. Daí somam-se as quantidades necessárias de RACKS e estima-se a potência necessária para alimentar a sala inteira. Depois somam-se as demandas das salas para determinar a potência da carga de TI. Finalmente, considerando o projeto das instalações, define-se a energia a ser adquirida da concessionária e dimensionam-se os equipamentos de energia (no-breaks, geradores e outros) e os equipamentos de refrigeração.

Os parâmetros relacionados à carga de TI, fundamentais no prosseguimento do projeto, devem ser definidos com base na criticidade, na capacidade e no crescimento, conforme ilustra a **Figura 9-4**.

- A criticidade define a tolerância aceitável para o downtime do negócio.
- A capacidade em KW tem a ver com a energia necessária para a carga de TI.
- O crescimento deve definir a necessidade de energia futura e deve descrever a energia necessária no início da operação, a energia mínima e máxima necessária após 3 anos, 5 anos, por exemplo. Os parâmetros de TI adotados devem estar alinhados aos requisitos de negócio.

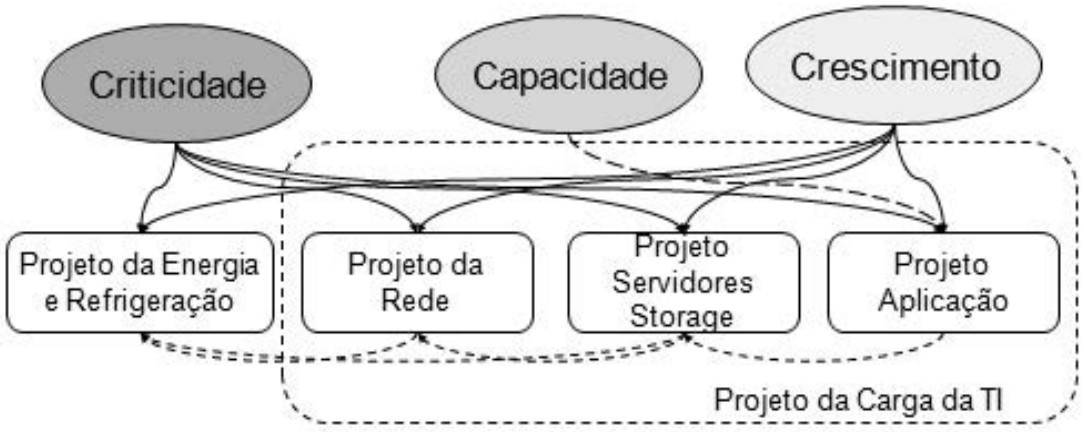


Figura 9-4 Parâmetros de TI no projeto do DATACENTER

O *Capacity Planning* é uma saída importante do Capacity Management. O Gartner afirma que, no caso do gerenciamento da capacidade de recursos do DATACENTER, existe grande dificuldade em se saber qual a real capacidade de energia e refrigeração requerida pelo DATACENTER e a instalação de novos equipamentos com alta densidade, como servidores *blades*, sem a devida consideração sobre a densidade de energia possível de ser realizada ser fonte de *downtime*, aquecimento e perda de redundância. Portanto, planejar a capacidade adequadamente é chave para o sucesso do projeto. Se a capacidade é planejada para baixo, pode inviabilizar o projeto já no curto prazo. Se planejada para cima, pode configurar um TCO excessivo para o projeto.

Para complicar, a carga de TI no cenário atual muda constantemente, os equipamentos mudam quase que diariamente e um ciclo de *refresh* de três anos obriga a organização a ter um bom gerenciamento da capacidade.

A APC sugere que o nível de controle do gerenciamento da capacidade tratado no DATACENTER seja o do RACK, e não no nível dispositivo (muito específico) ou da sala (muito abrangente).

É necessário preencher o DATACENTER com RACK de equipamentos segundo algum critério que categorize os *RACKS*. A maneira sugerida pela Sun é o que eles chamam de RLUs (*RACK location units*). A ideia que está por trás do uso das RLUs é flexibilizar o projeto do DATACENTER.

Qual a ideia, então? Toda a demanda de energia e refrigeração origina-se no RLU. A ideia é adotar um modelo hierárquico onde os RLUs definem as demandas de energia e refrigeração do DATACENTER. O fornecimento de energia deve ser sempre maior do que a demanda estimada pelas RLUs e vai depender do PUE da instalação para o DATACENTER. A **Figura 9-5** ilustra uma situação normal de projeto. A energia e a refrigeração devem atender às demandas das 105 RLUs.

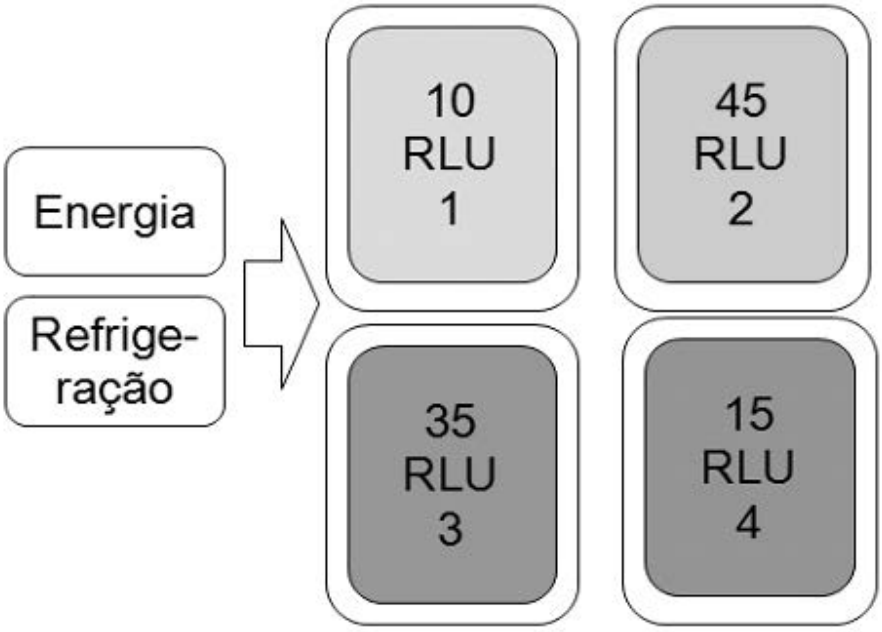


Figura 9-5 Energia e Refrigeração no DATACENTER

Os requisitos de uma RLU são:

- **Energia**: os aspectos de energia precisam ser todos considerados: como a energia chega ao RACK; o tipo de outlet; voltagem e amperagem, fase única ou trifásica; quanta energia o RACK demandará. Se o RACK tem energia redundante tem-se que considerar um fator (RM) no caso de 2.
- **Refrigeração**: os requisitos de refrigeração devem levar em consideração a eficiência dos dispositivos de refrigeração. A área necessária deve considerar a área do RACK mais a área de refrigeração.
- **Espaço físico**: dimensões do RACK. Dimensões da refrigeração do RACK. O espaço livre para os AISLES, no-breaks, rampas e área de livre circulação de ar.
- **Peso**: é crítico saber o peso de cada RACK e o peso total combinado dos RACKS. O projeto de um subfloor ou mesmo um raised floor deve considerar este peso.
- **Banda**: determina o cabeamento relativo a redes de storage e redes locais. A ideia aqui é dar a correta visão da necessidade de interfaceamento do RACK. Evidentemente, a banda fornecida pelos provedores deve ser equivalente ou exceder a banda necessária de cada RACK. Alterar a solicitação de banda é simples, mas alterar o cabeamento é complicado.
- **Capacidade funcional**: requerida somente para determinar a quantidade de RLUs e os tipos para atender o escopo do projeto. A capacidade funcional necessária (processamento, storage, etc.) é função da demanda das aplicações que rodarão no DATACENTER e fruto de outra discussão no capítulo 11.

A **Figura 9-6** ilustra uma RLU convencional e os atributos para a sua composição.

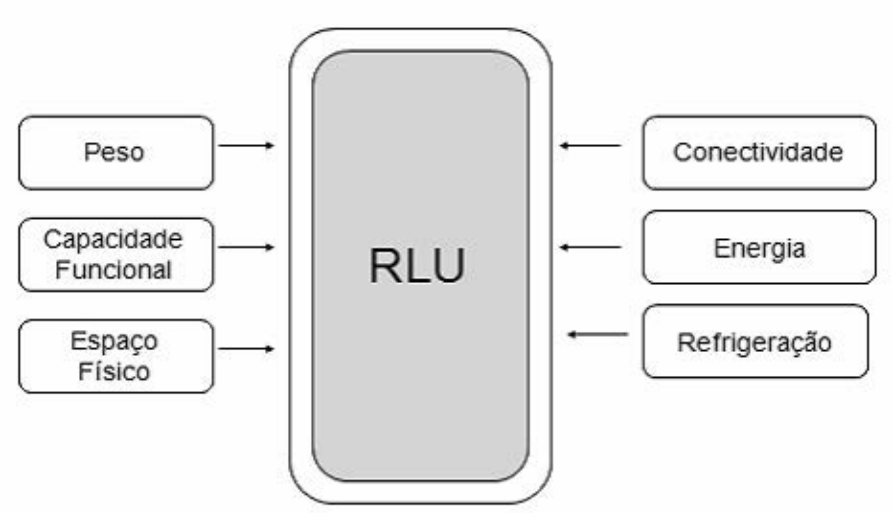


Figura 9-6 Dimensionamento baseado em RLUs

Com base nestas especificações, pode-se chegar aos requisitos de RLUs isolados ou em grupos e saber quantos RLUs o DATACENTER pode suportar. Esta é uma ideia realmente muito boa, pois, definindo os perfis (RLUs), é possível chegar aos grandes números necessários para o projeto do DATACENTER.

Os seguintes passos para determinar os componentes do DATACENTER são sugeridos:

- Determine quais equipamentos serão utilizados no DATACENTER com base no capacity planning.
- Defina os RLUs para os RACKS.
- Determine os 6 requisitos chave para as RLUs.
- Determine o número de RLUs necessárias para suportar o escopo do projeto.
- Determine os fatores limitantes.
- Defina o espaço final necessário.
- Defina os requisitos de energia e refrigerção.

9.2.2. Projeto da VIRTUALIZAÇÃO

Por sua vez, o projeto da VIRTUALIZAÇÃO pode ser dividido em duas partes ou em dois projetos.

O primeiro projeto é a justificativa em si para ir para a VIRTUALIZAÇÃO e a definição do fornecedor do software de VIRTUALIZAÇÃO. Não é uma tarefa fácil, pois os fabricantes dificultam a escolha, sempre publicando informações difíceis de serem comparadas, sempre focados em ampliar as vantagens do seu produto. Antes desta escolha é preciso aprovar o projeto dentro da organização, o que se conveciona chamar de *business case* ou estudo de viabilidade.

O segundo é relativo ao planejamento e à implementação da solução. Definir a arquitetura e a infraestrutura para suportar as aplicações no ambiente virtualizado é uma tarefa complexa, pois envolve diversas variáveis. Também se deve levar em consideração a necessidade de crescimento do ambiente e o consumo de recursos quando do uso de algumas funcionalidades. Em geral, tem-se duas alternativas:

- Se a infraestrutura a ser criada é nova, fica bem mais fácil, pode-se simular o ambiente a ser construído, construir um protótipo e depois partir para a solução.
- Se a infraestrutura já é existente, compartilhá-la com a VIRTUALIZAÇÃO requer um estudo sobre as suas reais possibilidades. Este é o foco maior deste capítulo, pois representa boa parte dos projetos existentes.

Os dois projetos são descritos na **Figura 9-7**.

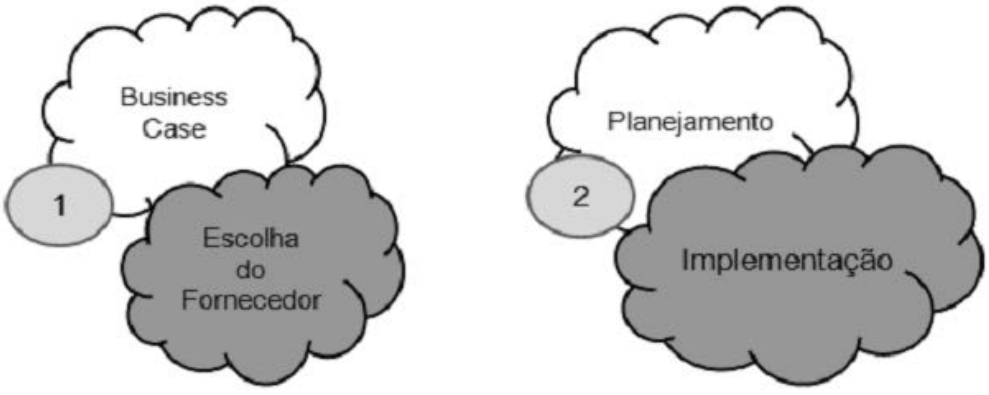


Figura 9-7 Projetos de VIRTUALIZAÇÃO

9.3. Projeto 1 – Business Case e Escolha do Fornecedor

O projeto 1 da VIRTUALIZAÇÃO inicia com o *business case*. O *business case* é a forma de comunicar para a alta direção os benefícios do novo projeto com vistas a aprová-lo. Ou seja, o primeiro passo de um projeto de VIRTUALIZAÇÃO é justificar o projeto para níveis superiores na organização. Depois, então, é feita a opção pelo fornecedor do software de VIRTUALIZAÇÃO e assim define-se a infraestrutura necessária, já uma etapa do Projeto 2.

As etapas iniciais do Projeto 1 de VIRTUALIZAÇÃO são ilustradas na **Figura 9-8.**

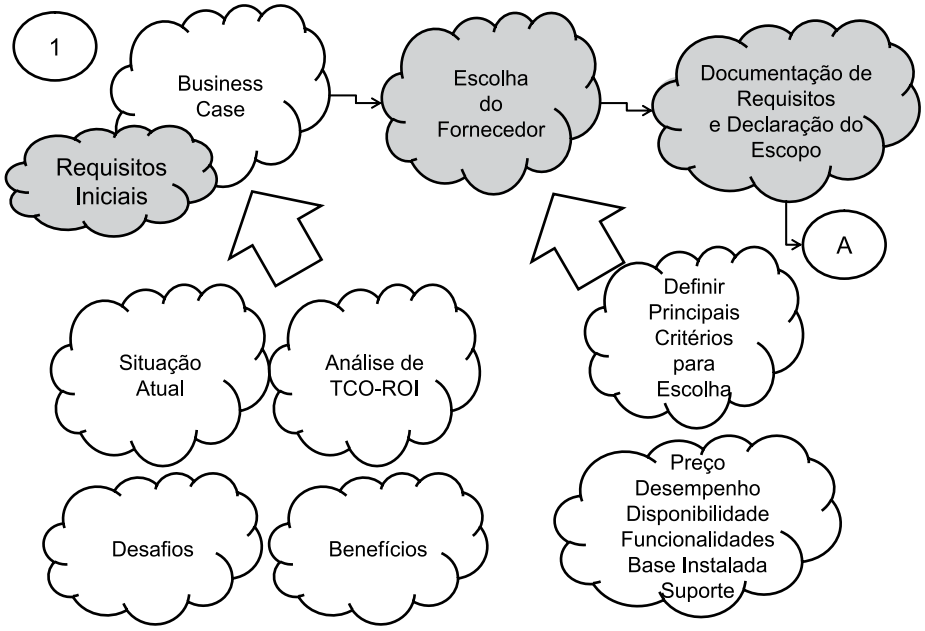


Figura 9-8 Projeto 1 de VIRTUALIZAÇÃO

9.4. Business Case

9.4.1. Introdução

Normalmente, em um *business case,* descrevem-se a situação atual e os principais desafios encontrados e apresenta-se uma proposta de solução com foco nos benefícios e lastreada em uma análise de TCO/ROI. Também faz-se um histórico de vida da infraestrutura de TI procurando posicionar a importância do projeto para o atual momento da organização e para o futuro. Já no business case, tem-se uma ideia de quais os requisitos iniciais do projeto serão utilizados na escolha do fornecedor.

O *business case* deve fornecer informações necessárias do ponto de vista do negócio para determinar se o processo justifica ou não o investimento.

Os indutores da VIRTUALIZAÇÃO são as novas demandas oriundas do CEO ou do CIO, do administrador de infraestrutura e tendências de mercado. Os CIOs priorizam usar a TI para o crescimento dos negócios, consolidar as operações, introduzir flexibilidade ao negócio e provar o valor do negócio de TI. Os administradores de infraestrutura querem reduzir o número de servidores, melhorar a disponibilidade dos sistemas, mitigar os riscos da mudança, reduzir o tempo para realizar o *deployment* de servidores e fazer mais com menos. O mercado se encarrega de dar os sinais de que uma nova tendência está chegando.

Essencialmente, é necessário entender os objetivos do negócio considerando que CEOs, CIOs e gerentes de infraestrutura possuem visões diferentes das necessidades, mas que convergem em um ponto quando o tema é transformar a situação atual da infraestrutura de TI da organização.

As justificativas para novos projetos que contemplem a VIRTUALIZAÇÃO passam por dificuldades atuais em provisionar novos servidores, dificuldades operacionais, consumo de energia, mau uso dos recursos, falta de estratégia para alta disponibilidade e recuperação de desastres, dificuldades com as aplicações legadas e novos hardwares.

9.4.2. Situação Atual

É preciso descrever a situação atual da infraestrutura de TI. Como poderia ser justificado um projeto de VIRTUALIZAÇÃO? É necessário descrever todas as dificuldades encontradas no gerenciamento e provisionamento de novos servidores, na utilização decorrente do pouco uso de recursos, ocupação de espaço físico, etc.

9.4.3. Desafios

Os desafios são diversos, entre eles alterar a infraestrutura existente em tempo real, que acontece na maioria dos casos, considerando que a organização continua a funcionar. O desafio maior é fazer a mudança para o novo ambiente de uma forma que haja o mínimo de paralisação das operações com a certa limitação de recursos humanos e físicos.

As organizações avançam na busca de soluções não proprietárias que simplifiquem o ambiente de infraestrutura de TI. Cerca de 90% dos servidores vendidos atualmente são baseados em plataforma x86, portanto, esta é uma tendência de mercado. A VIRTUALIZAÇÃO permite utilizar melhor os recursos de hardware, incluindo otimizar o uso dos recursos de processador e a memória dos servidores.

9.4.4. Benefícios

Novos negócios exigem uma infraestrutura de TI ágil e flexível para suportar a dinâmica dos processos de negócio. Este deve ser o foco quando da descrição dos benefícios do projeto. O uso da VIRTUALIZAÇÃO pode trazer grandes benefícios para a organização com a adequação da sua infraestrutura de TI ao negócio, mas requer planejamento e aquisição de novos recursos.

Os prováveis benefícios obtidos com a VIRTUALIZAÇÃO são:

- **Redução de custo total de propriedade (*TCO – Total Cost of Ownership*)**: o custo total de propriedade pode ser reduzido com o uso da técnica de VIRTUALIZAÇÃO. Os fabricantes disponibilizam ferramentas que permitem o cálculo do TCO, considerando a comparação de uma infraestrutura de TI com e sem VIRTUALIZAÇÃO. Em geral, é simples de justificar um projeto de VIRTUALIZAÇÃO utilizando a abordagem de TCO, tanto para a atualização da infraestrutura física já existente como também para a construção de uma nova infraestrutura. A redução do TCO é causada por:
  - **Redução do uso do espaço físico:** a utilização da VIRTUALIZAÇÃO permite a redução do espaço físico na medida em que considera a utilização de menos servidores como solução. Também a consolidação das estruturas de storage e backup, quase sempre contempladas num projeto de VIRTUALIZAÇÃO de servidores, acaba reduzindo a utilização do espaço como um todo.

- **Redução do consumo de energia**: quase sempre junto com a consolidação física vem a redução do consumo de energia. Servidores são os responsáveis pelo maior consumo de energia entre os equipamentos de TI e a consolidação acaba por reduzir o consumo de energia.
- **Isolação dos ambientes de testes, desenvolvimento e produção**: em muitas instalações, construir ambientes físicos diferentes para os ambientes de teste, desenvolvimento e produção pode ser muito caro. A utilização da VIRTUALIZAÇÃO otimiza o uso dos recursos, pois permite que estes ambientes existam de maneira completamente isolada, mesmo estando em poucos servidores físicos.
- **Flexibilidade na criação de novas máquinas virtuais**: as máquinas virtuais (servidores lógicos) podem ser criadas de forma automática em servidores já existentes. Na prática, a demanda por um novo servidor físico que dependeria de aprovação, compra, entrega, etc. pode ser atendida por uma máquina virtual pronta para rodar.
- **Padronização das plataformas**: na medida em que o HYPERVISOR passa a ser o elemento central do servidor virtualizado, todo o esforço de padronização de plataforma fica simplificado, pois a relação com o hardware se dá através dele. Diferentes sistemas operacionais podem coexistir sob a arbitração do HYPERVISOR.
- **Gerenciamento centralizado**: o gerenciamento das máquinas virtuais fica centralizado em uma única ferramenta, com única interface, reduzindo os custos operacionais de gerenciamento e promovendo a simplificação do ambiente.
- **Simplificação da implantação de técnicas de alta disponibilidade e recuperação de desastres**: a implantação de técnicas de alta disponibilidade, como clusters de servidores e o uso de tecnologias de replicação para suportar a recuperação a desastres, pode ser simplificada com o uso da VIRTUALIZAÇÃO. A VIRTUALIZAÇÃO permite a utilização do recurso de alta disponibilidade independentemente da técnica de cluster e facilita a criação do site secundário, otimizando os recursos alocados para o segundo site. Além disso, permite automatizar o processo de recuperação de desastres com a fácil integração promovida com técnicas de replicação do storage.
- **Viabilização da CLOUD COMPUTING e do DATACENTER DINÂMICO**: a VIRTUALIZAÇÃO é o componente central do datacenter dinâmico, que, por sua vez, viabiliza a CLOUD COMPUTING. A CLOUD COMPUTING e o datacenter dinâmico se viabilizam, na medida em que as soluções de VIRTUALIZAÇÃO avancem nos aspectos referentes a balanceamento de carga dinâmica, recuperação de falhas, segurança e interoperabilidade entre sistemas diferentes.

9.4.5. Análise de TCO/ROI

TCO

A metodologia de TCO considera quatro categorias de custo que contribuem para o custo: hardware e software, operações, downtime e administração. A metodologia do TCO considera os custos diretos e os custos indiretos como os custos de administração que são mais difíceis de serem medidos. O capítulo 1 abordou o TCO em maiores detalhes.

ROI

O ROI é uma medida usada para comparar o custo do projeto com os benefícios obtidos. A fórmula para calcular o ROI é dividir os benefícios realizados sobre um período de tempo dividido pela quantidade de investimento neste mesmo período.

ROI = Benefícios Quantificáveis/Custos Quantificáveis

Para avaliar um projeto, as organizações olham o ROI e o tempo requerido para que os benefícios ultrapassem os custos. No caso de projetos de VIRTUALIZAÇÃO, a análise de TCO é um pré-requisito para o estudo de ROI, desde que os clientes não consigam determinar a economia provocada pelo projeto sem saber qual o custo total de propriedade. O ROI para um projeto de VIRTUALIZAÇÃO é então calculado comparando o custo existente com a redução do custo e o custo evitado pelo novo projeto.

Os fabricantes de software, como a Microsoft e a VMware, possuem ferramentas exclusivas de cálculo de TCO produzidas pela Alinean, uma empresa que se originou do GARTNER. A VMware disponibiliza o VMware Virtualization TCO/ROI Calculator e a Microsoft disponibiliza o Microsoft Integrated Virtualization ROI Calculator. É possível readequar o projeto introduzindo na ferramenta de ROI os dados reais de inventário e de uso obtidos.

FERRAMENTAS DISPONIBILIZADAS PELA ALINEAN

VMWARE:

https://roianalyst.alinean.com/ent_02/AutoLogin.do?d=593411470991915416

MICROSOFT:

https://roianalyst.alinean.com/msft/AutoLogin.do?d=307025591178580657

A VMware disponibiliza em seu próprio site uma ferramenta mais simples para cálculo de TCO e ROI de projetos de VIRTUALIZAÇÃO baseados na sua plataforma. A ferramenta é chamada de VMware ROI TCO Calculator para Servidores e Desktop.

A **Figura 9-9** ilustra uma saída típica da ferramenta da VMware.

Cost Benefit Analysis		[Collapse]
	Over 3 yrs.	Over 5 yrs.
Investment	$223.651	$319.179
Savings		
Capital Expense	$196.050	$207.088
Operating Expense	$249.815	$434.175
Other	$18.152	$43.465
Total	$464.017	$684.728
ROI	107%	115%
Payback	1,6 years	1,8 years

Figura 9-9 VMware ROI TCO Calculator

9.5. Escolha do Fornecedor

Os critérios de escolha passam pelo desempenho do software de VIRTUALIZAÇÃO, disponibilidade, funcionalidades ofertadas, atual base instalada e suporte. A decisão hoje quase sempre recai entre as duas opções de fornecedores, que são a VMware e a Microsoft.

9.5.1. Custo e Desempenho

O estudo do Taneja Group, realizado em 2009, patrocinado pela VMware, é focado no parâmetro *VM Density* (densidade de máquinas virtuais). Segundo o estudo, o ESXi demonstrou nos testes uma vantagem da VMware em relação à Microsoft neste parâmetro de, no mínimo, 1.5:1. Ou seja, para uma dada carga de trabalho de aplicação, pode-se consolidar 8 servidores no Microsoft Hyper-V contra 12 servidores no VMware ESXi. Vale ressaltar que o estudo comparou as versões do VMware ESXi 3.5 com o Microsoft Hyper-V. A **Figura 9-10** ilustra a conclusão do trabalho.

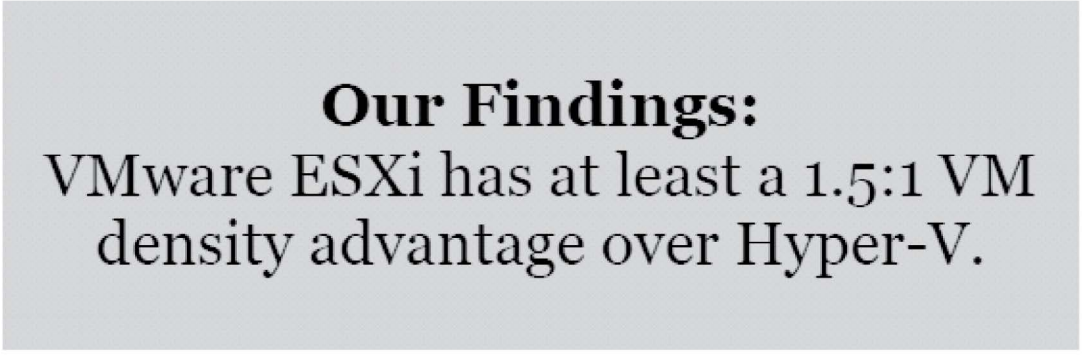


Figura 9-10 Resultado obtido pelo Taneja Group para VMware ESXi VS Hyper-V

Qual seria a importância desta diferença? Ela tem um impacto grande no custo da solução, na medida em que uma alta densidade de máquinas virtuais permite utilizar menos servidores, reduzir custos de licença dos SOs convidados e reduz o custo do gerenciamento, que em alguns casos tem a licença feita por host com HYPERVISOR.

Nas contas do Taneja Group, o custo por aplicação pode ser inclusive menor no VMware Infrastrucure (versão anterior ao vSphere), quando comparado ao Hyper-V e ao System Center.

A definição de densidade de VM é o número de VMs convidadas que possam ser efetivamente consolidadas em um HYPERVISOR sem impactos disruptivos no desempenho.

Os dois elementos que definem o custo de uma solução de VIRTUALIZAÇÃO são:

1. Eficiência do HYPERVISOR.
2. Eficiência da infraestrutura virtual de gerenciamento.

Estes dois elementos são responsáveis pelos custos, pois determinam a necessidade de:

1. Licenciamento.
2. Infraestrutura Física.

O Taneja utilizou dois testes para avaliar a densidade de VMs nos dois tipos de HYPERVISORS: DBHammer e SPECjbb. O paper conclui que o custo da ferramenta da VMware é 17% menor do que o custo da ferramenta da Microsoft.

A Microsoft, no paper *Virtualization Reality: Why Microsoft Virtualization Solutions Deliver Value when Compared to VMware,* justifica que no estudo Taneja não considerou o custo do VMware vCenter Server. Afirma também que implementações no mundo real não permitem as razões de densidade utilizadas para as máquinas virtuais encontradas nos testes. A Microsoft também critica a forma com que a VMware sugere adicionar mais servidores, quando os gargalos podem ser resolvidos adicionando memória, CPU e armazenamento, fazendo um aumento dos custos de forma incremental. Também a Microsoft alega que o pacote System Center permite obter mais funcionalidade do que o vSphere 4 e isto não é considerado. Segundo a Microsoft, a ferramenta de teste utilizada para simular o SQL Server não varia a carga e, portanto, o uso da memória no teste não é característica de um banco de dados real.

A Microsoft fornece em seu site uma ferramenta chamada *Virtualization Cost Comparison Calculator* para comparação de custos entre a sua plataforma e as versões do VMware vSphere. A **Figura 9-11** ilustra a tela da ferramenta.

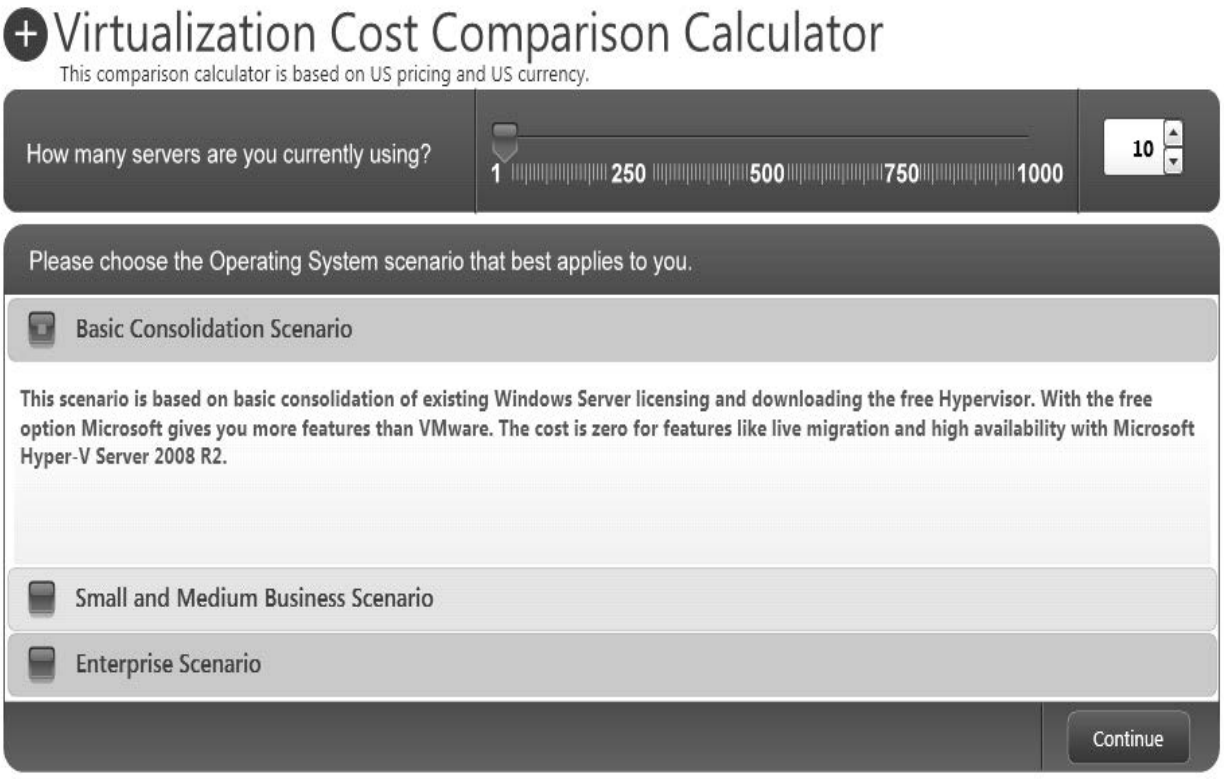

Virtualization Cost Comparison Calculator
This comparison calculator is based on US pricing and US currency.
How many servers are you currently using?
1
250
500
750
1000
10
Please choose the Operating System scenario that best applies to you.
Basic Consolidation Scenario
This scenario is based on basic consolidation of existing Windows Server licensing and downloading the free Hypervisor. With the free option Microsoft gives you more features than VMware. The cost is zero for features like live migration and high availability with Microsoft Hyper-V Server 2008 R2.
Small and Medium Business Scenario
Enterprise Scenario
Continue

Figura 9-11 MICROSOFT Virtualization Cost Comparison Calculator

A Microsoft possui dois conjuntos de opções para soluções de gerenciamento em ambientes virtualizados.

- SMSE (*System Center Management Suite ENTERPRISE*), utilizado para gerenciar ambientes com baixa densidade de máquinas virtuais.
- SMSD (*System Center Management Suite DATACENTER*), para instalações de alta densidade de máquinas virtuais.

As duas suites incluem licenças para:

- System Center Operations Manager 2007 R2.
- System Center Configuration Manager 2007 R2.
- System Center Data Protection Manager 2010.
- System Center Virtual Machine Manager 2008 R2.
- System Center Service Manager 2010.

A **Figura 9-12** mostra a comparação entre a suíte Microsoft SMSD com as opções da VMware para uma consolidação de 60 servidores físicos. Os preços são de lista para o mercado americano. No caso das suites SMSD e SMSE da Microsoft, é obrigatória a aquisição de *software assurance* para dois anos. A ferramenta da Micrososft permite

afirmar que o custo da solução da VMware é sete vezes maior do que o custo da ferramenta da Microsoft.

Figura 9-12 Custos para a VIRTUALIZAÇÃO de 60 servidores segundo a Microsoft

A VMware, por sua vez, publicou a comparação mostrada na **Tabela 9-1** no artigo *Why Choose VMWare*? Também neste estudo verifica-se a importância de considerar a densidade de máquinas virtuais nos cálculos do custo da solução.

1. VMware vSphere Enterprise Plus Edition.
2. VMware vSphere Advanced Edition.
3. VMware vSphere Standard Edition.
4. Windows Server 2008 R-2 (HYPER-V) e System Center.
5. Citrix XenServer 5.5 e Essentials Enterprise.

Tabela 9-1 Densidade de Máquina Virtual por Servidor Físico

	1	2	3	4	5
Número de Aplicações Virtualizadas	100	100	100	100	100
Número de VMs por Host	18	18	18	12	12
Número de Hosts	6	6	6	9	9
Custos de Infraestrutura (US$)	162.237	162.237	162.237	222.308	209,271
Custos de software (US$)	128.518	106.532	83.820	108.350	1211.993
Custos Totais (US$)	299.755	268.769	246.057	330.658	331.264
Custos por Aplicação (US$)	2.907	2.687	2.460	3.307	3.313

9.5.2. Funcionalidades

A comparação entre HYPERVISORS é mostrada na **Tabela 9-2** para:

1. VMware ESX 4.1.
2. Microsoft Windows Server 2008 R-2 com HYPER-V.
3. Citrix XenServer 5.5.

Tabela 9-2 Comparação entre HYPERVISORS

ATRIBUTOS	1	2	3
Espaço em disco	70MB	3GB com instalação Server CORE	1.8GB
Independência do SO	Não é baseado em tipo específico de SO	Baseado no WS08 na partição pai	Baseado em Linux e partição Dom0
Drivers	Otimizado com os fornecedores de hardware	Drivers Windows genéricos	Drivers Linux genéricos
Gerenciamento da memória	Consegue reaproveitar memória não utilizada	Não consegue reaproveitar memória não utilizada	Consegue reaproveitar memória não utilizada, mas não ajusta a alocação de memória baseado no uso por VM
Gerenciamento do armazenamento	VMware vStorage VMFS, Storage vMotion	Não faz migração do storage	Não faz migração do storage. Funcionalidades de storage suportam poucos arrays
Escalabilidade do I/O	Modelo Direct Driver	Gargalo de I/O na partição pai do SO	Gargalo de I/O na partição Dom0 do SO
Gerenciamento de recursos do Host	Faz qualidade de serviço para I/O de storage e rede		
Desempenho	*Large memory pages*, 8-way vSMP, paravirtualização VMI, VM-directpath I/O, PV guest SCSI driver	*Large memory pages*, 4-way vSMP só no WS08 e VMs Windows 7	*Large memory pages*
Tecnologia Virtual Security	VMware VMsafe security API		
Alocação de Recursos Flexíveis	Adição quente de VM vCPUS e memória, crescimento do volume VMFS, extensão de discos virtuais a quente	Extensão de discos virtuais a quente	

A comparação dos serviços Live Migration, Capacidade de Agregação e Alocação de Recursos e a Capacidade de Gerenciamento é mostrada nas **Tabela 9-3**, **Tabela 9-4** e **Tabela 9-5,** onde:

1. VMware vSphere 4.1 com vCenter Server.
2. Microsoft Server 2008 R-2 com Hyper-V e SCVMM-R2.
3. Citrix XenServer 5.6 com Essentials Enterprise Edition.

Tabela 9-3 Comparação dos Serviços Live Migration

FUNCIONALIDADES	1	2	3
Downtime Zero para a aplicação	OK	OK	OK
Migração de múltiplas VMs concorrentemente	OK	Só uma migração por HOST ao mesmo tempo	Só uma migração por HOST ao mesmo tempo
Manutenção dos parâmetros do switch virtual e dos parâmetros de segurança da rede durante a migração viva de um HOST para outro	OK		
Migração viva do storage	OK		
Migração viva de um protocolo de storage para outro	OK		

Tabela 9-4 Comparação das Capacidades de Agregação e Alocação de Recursos

FUNCIONALIDADES	1	2	2
Pools de Recursos Hierárquicos	OK		
Isolação entre Pools de Recursos	OK		
Gerenciar um ÚNICO switch virtual para um cluster inteiro	OK		
Forçar os parâmetros de segurança de rede no nível da VM	OK		
Affinity Rules	OK		
Habilidade de fazer por VM a banda de I/O do storage	OK		
Habilidade de fazer por VM a banda de I/O do storage	OK		
Utilizar Live Migration para downtime zero e load balancing automático	OK	PRO Tips requer Systems Center Operations Manager, não propicia heurística	Citrix WLB, não propicia heurística

Tabela 9-5 Comparação das Funcionalidades de Gerenciamento

FUNCIONALIDADES	1	2	3
Hot VM Cloning	OK		
Provisão de VM templates	OK	OK	OK
Customização automatizada do convidado	OK	Só com Windows	
Configuração de Host centralizada	OK		OK
Client web para gerenciamento	OK		
Controle avançado de recursos de CPU	OK	OK	
Controle avançado da banda de rede	OK		OK
Alarmes customizados	OK	Com SCOM	OK
Agendamento de tarefas	OK		
Reports customizados	OK	Com SCOM	
Mapas da topologia dos recursos	OK		
Tracking de tarefas e eventos	OK		
Escolha do banco de dados de gerenciamento	OK		
Abstração do gerenciamento em alto nível com host profiles e vApps	OK		
Client único para conexão a plataforma de gerenciamento centralizada & HOST standalone	OK		

A **Tabela 9-6** compara as funcionalidades para recuperação a desastres:

Tabela 9-6 Comparação da Capacidade de Recuperação a Desastres

Funcionalidades	1	2	3
Integração do software de virtualização com a replicação do storage	OK		StorageLink com pouco HCL
Criação gráfica de planos de recuperação	OK		
Teste não disruptivo One-Button a qualquer tempo	OK		
Automação One-Button do failover do DR	OK		
Histórico detalhado de teste e recuperação do DR	OK		

A **Tabela 9-7** compara as funcionalidades de gerenciamento de patches e update.

Tabela 9-7 Comparação de Funcionalidades de Gerenciamento de Patches e Update

FUNCIONALIDADES	1	2	3
Wizard Patch Management Integrado	OK		
Instalado como um plug-in para a interface de gerenciamento	OK		
Patching offline seguro de VMs	OK		
Patching de HOST com zero-downtime	OK	OK	OK
Cria múltiplos baselines	OK		
Download automático de Patches	OK	OK	
Snapshots automatizados com prioridade sobre o patching	OK		

A **Tabela 9-8** compara o hardware suportado.

Tabela 9-8 Comparação baseada em Hardware Suportado

	VMware ESX 4.0	CITRIX XENSERVER 5.5
Servidores suportados	>1550	244
Controladoras de storage suportadas	>650	112
Storage suportado	>850	113
Cartões de rede suportados	>350	105

A **Tabela 9-9** compara a quantidade de sistemas operacionais suportados.

Tabela 9-9 Comparação sobre a quantidade de SOs Convidados

SO convidado suportado	VMware ESX 4.1	MS WS08 R2 com HYPER-V	CITRIX XEN SERVER 5.5
Total	65	17	25

9.5.3. Base Instalada

A base instalada conta. Um software com grande base instalada normalmente tem um processo de correção de bugs e fornecimento de patches aprimorado. Também as redes que se formam e os fóruns possibilitam rapidamente resolver uma questão técnica.

9.5.4. Suporte

Suporte trata de saber como funciona a rede do fornecedor de software no que diz respeito a suporte. Suporte em países que utilizam um idioma diferente do inglês é um aspecto relevante. Também o SLA dos serviços baseados em 0800 e Internet devem ser verificados.

9.6. Projeto 2 – Solução de VIRTUALIZAÇÃO

9.7. Iniciação

Escolhido o fornecedor, as funcionalidades a serem utilizadas definem a edição do software a ser adquirida. Depois é necessário definir a arquitetura e a infraestrutura para suportar o novo projeto.

9.8. Planejamento

A **Figura 9-13** ilustra o fluxo inicial do projeto 2.

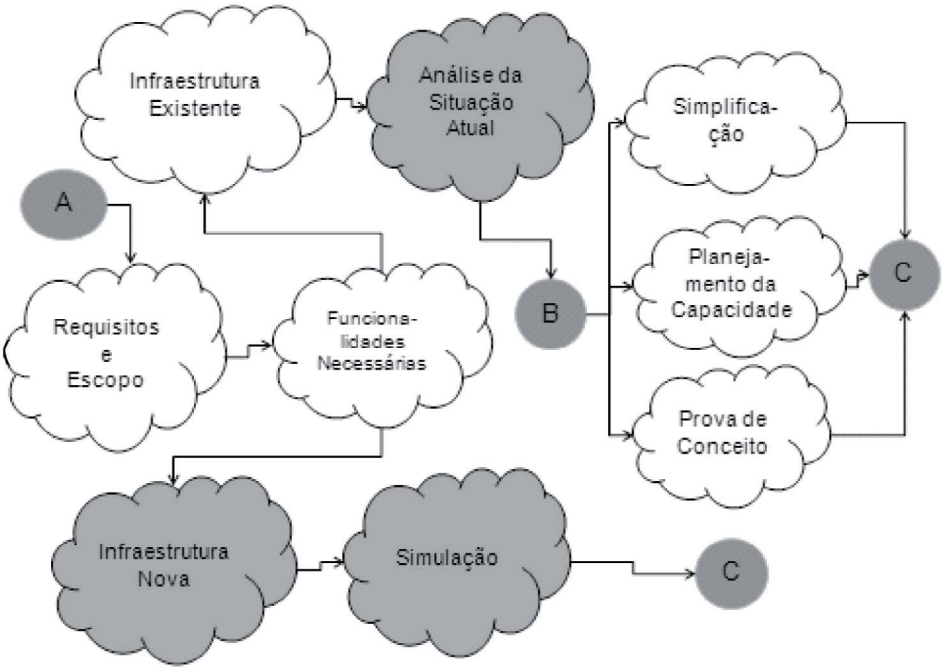


Figura 9-13 Projeto 2 de VIRTUALIZAÇÃO – Etapa 1

9.8.1. Documentação de Requisitos

A documentação de requisitos deve conter como os requisitos iniciais atendem o projeto.

9.8.2. Declaração do Escopo

A declaração do escopo deverá conter:

- Descrição do escopo.
- Critérios de aceitação do produto.
- Entregas do projeto.
- Exclusões do projeto.
- Restrições do projeto.

9.8.3. Criação da EAP

Normalmente em um projeto de VIRTUALIZAÇÃO define-se uma Estrutura Analítica do Projeto (EAP) com fases, etapas e grandes atividades. A EAP deve representar o escopo.

9.8.4. Atividades em Termos de Tempo, Custo e Qualidade

Atividades com cronograma, recursos e entregas. Os pacotes de trabalho ou atividades resumo devem ser detalhados até onde interessa ao gerente de projeto.

9.8.5. Identificação dos Riscos

A identificação dos riscos é uma etapa básica do projeto. É importante documentar os riscos identificados e definir a estratégia de resposta aos riscos.

9.8.6. Análise da Situação Atual

A infraestrutura pode já existir e, neste caso, deverá passar por grandes ajustes ou o projeto pode ser novo. No caso do projeto novo, será necessário simular as cargas das aplicações de TI no ambiente virtualizado.

- Se a infraestrutura já existe, caso mais comum, faz-se um levantamento da situação de carga encontrada para a definição da infraestrutura virtualizada. Os fabricantes como a VMware e Microsoft disponibilizam ferramentas que permitem avaliar a situação da infraestrutura existente em termos de processamento, memória, I/O e rede dos servidores.

Deve-se considerar a análise da situação atual no caso de um projeto para uma infraestrutura existente. As funcionalidades escolhidas para o projeto de VIRTUALIZAÇÃO definirão os recursos a serem aproveitados da antiga infraestrutura. Deve-se definir uma meta de utilização para os processadores lembrando que o uso de determinados recursos, como a alta disponibilidade e o balanceamento dinâmico das cargas de trabalho, obrigam a nova infraestrutura a ter uma folga, tanto nos recursos de processamento como nos recursos de memória. O ideal é trabalhar a infraestrutura nova para que ela funcione a média carga (em torno de 50%).

Informações importantes a serem levantadas sobre a infraestrutura existente são:

- Número de servidores.
- Processamento total dos servidores.
- Memória total dos servidores.
- Utilização de banda por servidor.
- Espaço total de disco dos servidores.
- IOPS de disco médio por servidor.
- Se a infraestrutura não existe deve-se utilizar um software de simulação com uma situação real de carga para definição da infraestrutura. Esta opção em geral é complexa e as ferramentas existentes são proprietárias.

No caso de uma infraestrutura nova, é necessário definir o perfil das cargas de trabalho para os diversos servidores, incluindo:

- Servidores de Teste e Desenvolvimento.

- Servidores de Arquivo e Impressão.
- Servidores web.
- Servidores Controladores de Domínio.
- Servidores de DNS.
- Servidores Genéricos.

9.8.7. Considerações sobre a DMZ

Considera-se prudente em projetos de VIRTUALIZAÇÃO deixar servidores de Firewall e servidores da DMZ separados dos servidores de VIRTUALIZAÇÃO usados na rede interna do DATACENTER. Do ponto de vista técnico, fazer a DMZ com servidores virtualizados funciona, mas qualquer descuido do administrador com configurações lógicas pode abrir uma porta para invasão por hackers.

9.8.8. Considerações sobre a VIRTUALIZAÇÃO das Principais Aplicações

O planejamento da capacidade (*capacity planning*) visa descrever as necessidades de capacidade presentes e futuras da infraestrutura de TI em função das mudanças esperadas na demanda por serviços de TI, da substituição de componentes obsoletos e das novidades tecnológicas. O plano gerado deve definir as mudanças necessárias para o atingimento dos níveis de serviço como também os custos associados à mudança.

O processo de *capacity planning* é a grande dificuldade de qualquer projeto de infraestrutura, devido à incerteza natural para a realização deste tipo de planejamento. Existem muitos elementos envolvidos em um projeto de capacidade e muitas das informações necessárias para a realização do correto planejamento não estão disponíveis. O projetista então está sempre entre dois extremos: se avalia o risco para cima, utiliza um hardware de maior desempenho e armazenamento, aumenta o custo. Se avaliar o risco para baixo, pode definir um hardware de menor desempenho e condições de armazenamento a um custo mais aceitável, mas pode ter problemas e aumentar o risco. A linha de corte entre estes extremos é sempre complexa e o bom projeto deve buscar um equilíbrio entre estes extremos.

A alteração do cenário de infraestrutura saindo de computadores de grande porte (*mainframes*) para DATACENTERS baseados em ambiente distribuído, provocou também a troca da disciplina de *capacity planning* pela técnica de provisionar para cima os servidores. A adição de caixas (boxes) extras ou de caixas de maior processamento serve para compensar a falta de método na determinação do correto hardware. Esta maneira de provisionar foi amparada pelos preços agora muito abaixo das soluções baseadas em *mainframe*.

Esta técnica com a tendência de consolidação dos DATACENTERS está se mostrando inadequada, pois a organização passa a depender completamente deste único ponto de provimento de processamento e armazenagem, e um erro no dimensionamento pode trazer sérios transtornos à organização.

Assim, diversas técnicas surgiram com o objetivo de melhorar o processo de planejamento da capacidade. Hoje, existem basicamente três técnicas para a realização do *capacity planning*[12]:

- ***Capacity Benchmarking***: O *capacity benchmarking* (também conhecido como *LOAD TESTING*) é a técnica mais utilizada e normalmente a mais cara. É necessário construir um sistema, fazer o setup da configuração e submeter o tráfego para ver como o sistema se comporta. O resultado do benchmarking pode então ser utilizado para a comparação de desempenho entre sistemas distintos.
- ***Capacity Trending:*** O *capacity trending* é realizado com análises lineares ou estatísticas. Normalmente esta técnica aponta uma saída para o problema, mas existe dificuldade com a otimização dos resultados.
- ***Capacity Modeling:*** O *capacity modeling* é uma técnica que utiliza simulação ou modelamento analítico e que permite testar várias soluções para um problema sem precisar necessariamente que sejam implementados.

O *capacity planning* é necessário, na medida em que existe uma relação entre as necessidades do negócio e a infraestrutura de TI. O *capacity planning* está associado à possibilidade de melhor projetar a infraestrutura de TI, fugindo da estratégia do servidor super provisionado comentada anteriormente. A ideia do *capacity planning é* realizar um projeto para a correta WORKLOAD a que será submetido o servidor e sua tendência de uso futura. Realizar o *capacity planning* dos servidores e da rede são disciplinas importantes para permitir a construção de um DATACENTER eficiente.

A **Figura 9-14** ilustra o efeito do HYPERVISOR no *capacity planning* das aplicações. É evidente que a camada do HYPERVISOR introduz uma perda de desempenho, considerando que é mais uma camada entre as aplicações e o hardware.

Como a VIRTUALIZAÇÃO consiste basicamente em se pôr uma camada de software a mais em um sistema computacional, a questão sobre o quanto isso afeta o desempenho final é imediata. Estudos feitos pela VMware e pela XenSource apontam para um queda de desempenho causada pela presença do HYPERVISOR, em geral, entre 2% e 10%, com algumas situações impondo perdas maiores. Cabe ressaltar que esses resultados foram obtidos usando benchmarks genéricos.

As perdas são pequenas porque os fabricantes se uniram em torno da ideia de viabilizar a VIRTUALIZAÇÃO em baixa plataforma e introduziram diversas modificações nas várias camadas que suportam a VIRTUALIZAÇÃO, com especial ênfase na melhoria do desempenho do processador e da memória em servidores. Além disso, as funcionalidades incorporadas trazem outros benefícios que mais do que compensam as perdas de desempenho introduzidas pelo HYPERVISOR.

12 Ver Capacity Planning: Discipline for DATA CENTER Decisions. TeamQuest, 2004.

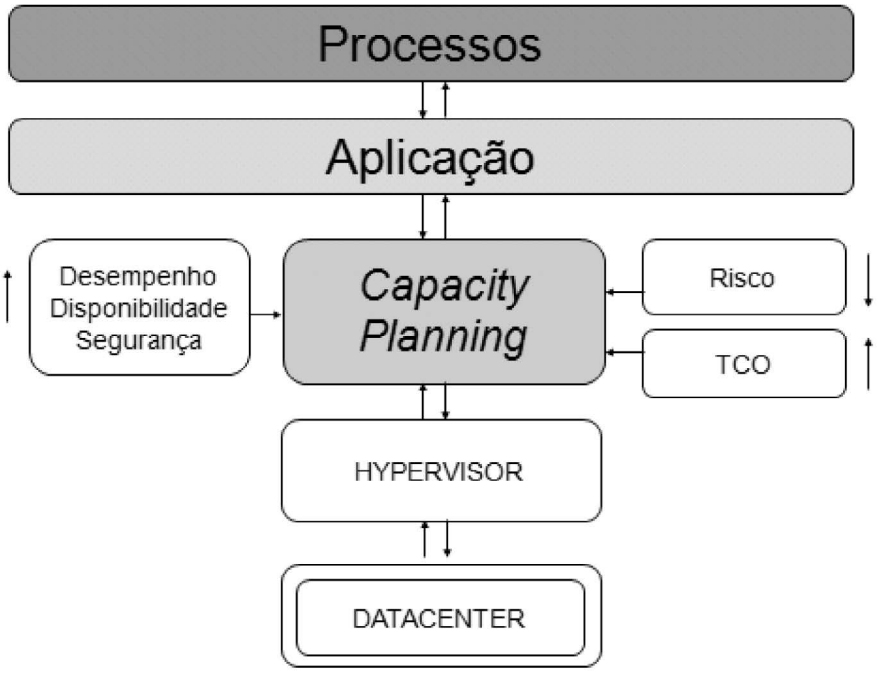


Figura 9-14 Capacity Planning com VIRTUALIZAÇÃO

Um *capacity planning* tem basicamente três objetivos:

- Determinar os níveis de serviço necessários.
- Analisar a capacidade corrente.
- Planejar a infraestrutura para o futuro.

9.8.9. Benchmarking da Indústria de TI

O processo de comparação do desempenho entre dois ou mais sistemas é chamado de *benchmarking*. Os *benchmarkings* são a maneira que a indústria de TI encontrou de permitir a comparação de desempenho fornecido por diferentes arquiteturas tecnológicas para uma mesma aplicação.

Com a evolução das arquiteturas utilizadas pelos servidores, fica cada vez mais difícil comparar o desempenho das aplicações rodando em diferentes arquiteturas somente olhando as especificações dos componentes. Um exemplo é o caso específico dos processadores que utilizam arquiteturas diferentes. Levar em consideração só a velocidade do processador para comparar processadores Intel e AMD, por exemplo, ou comparar a arquitetura x86 com a arquitetura RISC é inadequado com a variedade de parâmetros envolvidos (pode-se citar o cache dos processadores, a arquitetura da memória, tipo de compilador utilizado, etc.) que influenciam o desempenho.

Mesmo a comparação entre arquiteturas x86 Intel de processadores que utilizam diferentes números de núcleos (cores) é difícil, quando observa-se só especificações técnicas como a velocidade em GHz dos processadores.

Por isso, testes foram desenvolvidos para serem realizados entre diferentes arquiteturas, permitindo que esses resultados possam ser comparados por aplicação. Ou seja, realiza-se um mesmo teste, que considera as características de uma determinada aplicação em arquiteturas diferentes e, assim, consegue-se ter uma referência de desempenho.

Existem diversos *benchmarkings* na indústria de TI e o usuário da aplicação deve utilizar um benchmarking adequado para obter um quadro real do desempenho da aplicação.

A **Figura 9-15** ilustra os principais benchmarkings utilizados pela indústria de TI por aplicação. Associou-se o *benchmark* ao tipo de aplicação com duas tonalidades, uma mais escura e outra mais clara. Se a cor for escura, entende-se que o *benchmark* é adequado, se a cor for mais clara, entende-se que o *benchmark* pode ser aplicável, mas com algumas limitações. Os *benchmarkings* relacionados podem ser utilizados para os seguintes tipos de aplicações:

- 1=banco de dados.
- 2=banco de dados para suporte a decisão.
- 3=correio eletrônico.
- 4=servidor de aplicação.
- 5=servidor web.
- 6=aplicações JAVA VM.
- 7=servidor de arquivos.
- 8=virtualização.

Importante salientar que existem alguns outros *benchmarkings* que estão associados a subsistemas específicos como, por exemplo, o *IO Meter* associado ao subsistema de disco (I/0) e *benchmarkings* para aplicações específicas, como o LINPACK para *cluster*s HPC. Também recentemente foi lançado o TPC-E, que promete ser mais preciso para verificação de bancos de dados OLTP.

TPC-C

Simula um ambiente de computação onde usuários executam transações contra um banco de dados. A *performance* do TPC-C é medida em novas transações por minuto. As principais métricas são: taxa de transação (tpmC) e preço por transação ($/tpmC). Quanto maior o tpmC melhor. Quanto mais baixa a relação $/tpmC melhor.

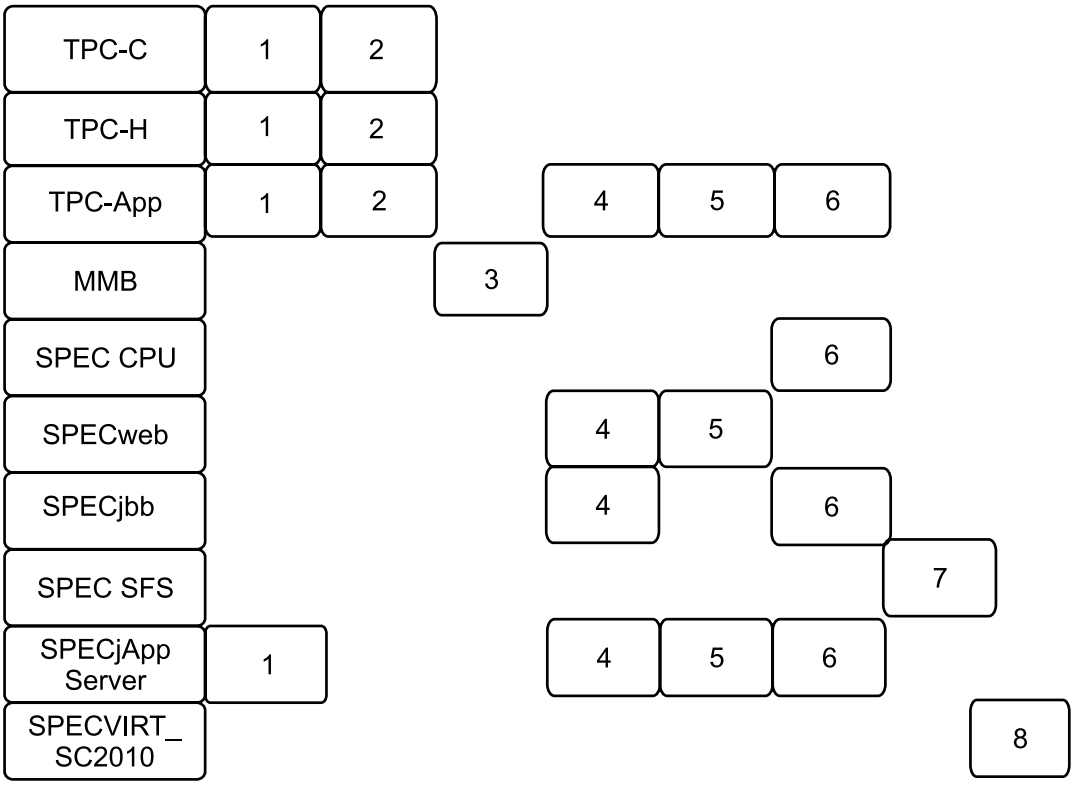


Figura 9-15 Benchmarkings da Indústria de TI

TPC-H

É um benchmark para sistemas de apoio à decisão. Consiste em um conjunto de *queries ad-hoc* orientadas a negócio e considera também modificações de dados concorrentes. As métricas são *composite-query-per-hour performance* (QphH) e preço/performance ($/QphH).

TPC-App

É um benchmark para servidor de aplicação e *web services*. A carga de trabalho (*WORKLOAD*) é realizada em um ambiente que simula transações *web business-to-business* em um servidor de aplicação que opera 24 por 7. As principais métricas são *web services interaction per second* (SIPS) por servidor de aplicação e preço/performance ($/SIP).

MMB

O MMB (*mapi messaging benchmark)* mede a performance e escalabilidade de servidores rodando o Microsoft Exchange Server. Este benchmark foi desenvolvido para ajudar a definir a configuração do servidor de Exchange para usuários típicos de correio/mensagens. O MMB esta relacionado à versão do Exchange – por exemplo, MMB3 está relacionado com o MS Exchange Server 2003.

SPEC CPU

A organização SPEC (file://www.spec.org/) foi criada para estabelecer, manter e endossar um conjunto padrão de benckmarks úteis na definição e na criação de uma linha de base (*baseline*) para medir a performance de servidores. Especificamente, o benchmark CPU (CPU2006) permite medir a performance de sistemas que envolvem CPU, arquitetura de memória e compilador. As métricas são SPECint2006 e SPECint_rate2006, que permitem comparar a velocidade e a performance para números inteiros (do ponto de vista de *throughput* de processamento) entre servidores distintos, e SPECfp2006 e SPECfp_rate2006, que permitem comparar a velocidade e a performance de servidores para operações de ponto flutuante. O benchmark SPEC CPU é relevante para a maioria das aplicações que processam números, mesmo não sendo específico para nenhuma aplicação.

SPC

O SPC não faz parte do diagrama mostrado anteriormente, mas é um *benchmarking* muito utilizado para comparar diferentes sistemas de armazenamento. O *benchmark* SPC propicia uma fonte de comparação para o desempenho de dispositivos de armazenamento. Os *benchmarks* SPC são projetados por fornecedores independentes e são aplicáveis em um largo número de unidades de armazenamento com configurações e topologias diversas. Existem quatro SPC *benchmarks*: SPC Benchmark 1 (SPC-1), SPC Benchmark 2 (SPC-2), SPC Benchmark 1C (SPC-1C) e SPC Benchmark 2C (SPC-2C). O SPC-1 e o SPC-1C são focados em operações aleatórias de I/O como OLTP. O SPC-2 e o SPC-2C são focados em operações de I/O com acesso de dados sequenciais de larga escala como *video on demand*.

SPECVIRT_SC2010

O SPECvirt_sc2010 é o primeiro *benchmark* focado na avaliação de servidores de DATACENTER utilizados em um cenário de consolidação. O SPECvirt_sc2010 mede o desempenho do sistema fim a fim, incluindo todos os componentes como hardware, plataforma de virtualização, sistema operacional convidado e aplicações. O *benchmark* pode ser utilizado para comparar a virtualização por hardware, a virtualização pelo sistema operacional e esquemas baseados em particionamento do hardware.

O benchmark SPECvirt_sc2010 utiliza várias WORKLOADS representando aplicações que são comuns em cenários de VIRTUALIZAÇÃO. Estas WORKLOADS são modificadas para a obtenção de um típico cenário de consolidação para CPU, memória, disco e rede para cada WORKLOAD. Estas WORKLOADS são versões modificadas de SPECweb2005, de SPECjAppServer2004 e de SPECmail2008. A escalabilidade é conseguida rodando conjuntos adicionais de máquinas virtuais até que o throughput total alcance um pico. Todas as VMs devem continuar a atender a qualidade requerida de QoS.

O *benchmark* SPECvirt_sc2010 também inclui opções para medir a energia consumida e a relação energia/desempenho. Baseado na metodologia SPECpower, a medição de energia requer o uso de analisadores de potência adequados. O *benchmark* tem

opção de rodar com a monitoração de energia habilitada e pode fornecer resultados em três categorias:

- Desempenho somente (SPECvirt_sc2010).
- Desempenho/WATT para o sistema sobre teste, incluindo servidor e storage (SPECvirt_sc2010_PPW).
- Desempenho/WATT só para servidor SPECvirt_sc2010_ServerPPW).

O benchmark pode ser utilizado por fornecedores de hardware, fornecedores de software e fornecedores de aplicação, além de pesquisadores.

9.8.10. Opções para Definir a Infraestrutura

Pode-se dividir um ambiente típico de *capacity planning* para VIRTUALIZAÇÃO de infraestrutura em três níveis, conforme ilustrou a **Figura 9-13**.

Simplificação

1 a 20 servidores: para efeito de um projeto de VIRTUALIZAÇÃO, normalmente o simples planilhamento das informações sobre os servidores existentes é suficiente para fazer um projeto. As informações necessárias são: servidores existentes, configurações incluindo versão do sistema operacional e aplicação, consumo de processamento, memória, disco e rede. Também é importante o possível crescimento da base de dados e detalhes do tipo de backup utilizado.

Capacity Planning

20 servidores a 60 servidores: sugere-se a utilização de uma ferramenta do tipo VMware Capacity Planner Tool por pelo menos quinze dias. O cliente precisa disponibilizar um servidor com o sistema operacional Windows Server 2003 standard ou superior, uma lista com os servidores alvos da VIRTUALIZAÇÃO e portas apropriadas, acesso ao domínio e uma conexão VPN. O *VMware Capacity Planner Tool* permite simplificar o processo de avaliação e fornecer dados que permitem definir o cenário ideal para a consolidação. O *VMware Capacity Planner Tool* também permite modelar cenários de consolidação.

O planejamento da capacidade de projetos de VIRTUALIZAÇÃO é um dos principais aspectos a serem considerados. Os fabricantes disponibilizam ferramentas que permitem o correto planejamento da capacidade baseado na utilização da infraestrutura existente. Planejar a infraestrutura consolidada necessária para um projeto de VIRTUALIZAÇÃO, considerando a existência de diversos servidores que atendem a diversas WORKLOADS provenientes de diferentes departamentos, é uma tarefa árdua. Também as métricas utilizadas precisam ser cuidadosamente consideradas.

Os fabricantes de softwares como a VMware, entendendo esta dificuldade, disponibilizaram rapidamente ferramentas que permitem fazer o inventário do ambiente de TI incluindo servidores, storage e rede. Seu real uso é fornecer sugestões para o novo ambiente virtualizado que otimizam a nova infraestrutura. As ferramentas de planejamento da capacidade combinam o inventário com as informações de desempenho. O uso dos

recursos em ambientes não virtualizados em geral é muito baixo, o que permite realizar projetos de VIRTUALIZAÇÃO com altas taxas de conversão. Também é necessário considerar a otimização do uso dos recursos baseada no tipo de recurso utilizado. A **Figura 9-16** (a) e (b) ilustra o caso.

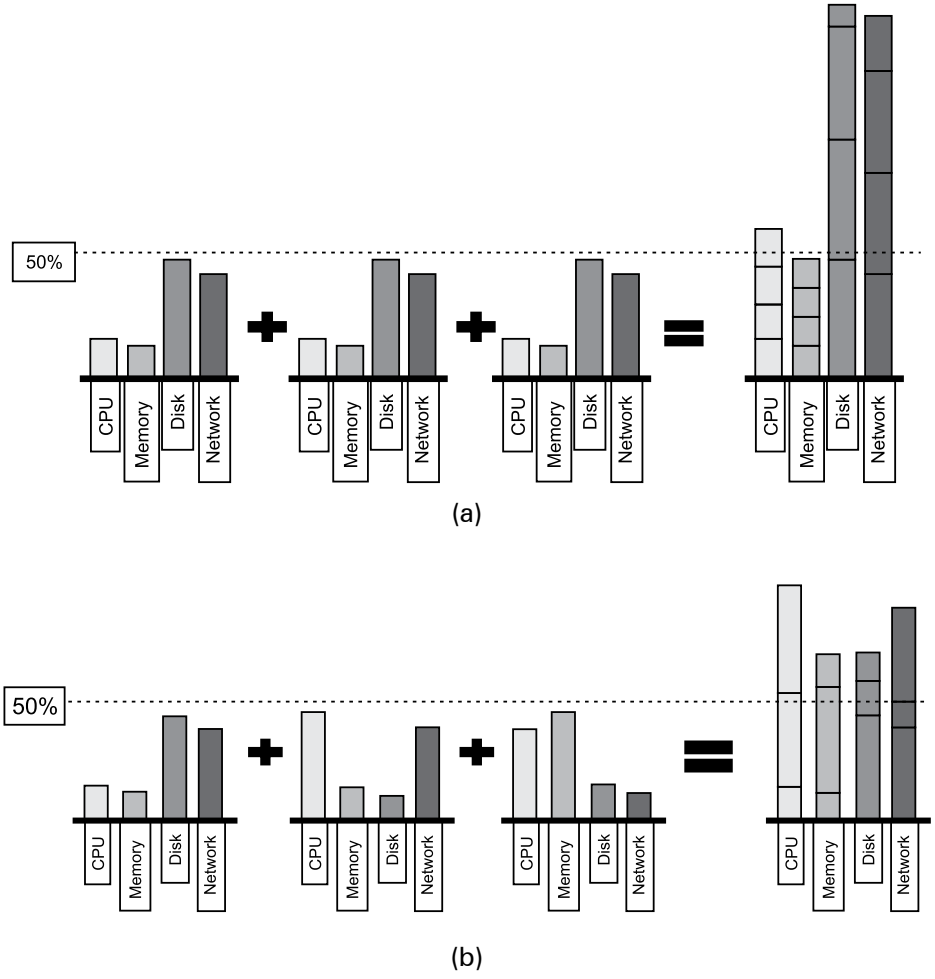


Figura 9-16 Otimização baseada no uso de recurso (a) RUIM (b) MELHOR

VMware Capacity Planner

O VMware Capacity Planner é uma ferramenta baseada em interface Web que consolida todas informações de projetos em uma grande base de informação que serve para o fornecimento de benchmarks para a indústria. Os principais componentes do VMware Capacity Planner são o Data Collector, Data Manager, Data Analyzer, o Information Warehouse e o Dashboard.

- ***Data Collector*** **(Coletor de dados):** instalado localmente no site do cliente, ele utiliza um agente para descoberta de sistemas, recolher um inventário detalhado de hardware e software e também coletar métricas de desempenho, que são exigidas para análise. O coletor pode reunir dados a partir de múltiplos ambientes (Windows e Linux).

- ***Data Manager* (Gerenciador de Dados):** este componente gerencia os dados recolhidos durante o processo de coleta, fornecendo uma visão organizada das informações coletadas. O Data Manager também envia os dados coletados das informações para o Warehouse da VMware.
- ***Information Warehouse:*** uma central de dados hospedada na VMware onde os dados coletados pelo Capacity Planner instalado no site serão agregados, organizados e preparados para análise. Na Warehouse, criamos estimativas para obter o cenário mais próximo do ideal possível.
- ***Data Analyzer:*** este componente funciona como o núcleo analítico que processa todas as análises exigidas para um planejamento inteligente. Possui avançados algoritmos com a capacidade de resolver problemas de otimização.
- ***Dashboard:*** um aplicativo baseado na Web que oferece acesso às informações parciais ou totais do andamento da coleta através da interface do browser. Os usuários podem acessar remotamente um rico conjunto de cenários de análises pré-construído e criar relatórios personalizados a partir dos dados coletados, podendo fazer também testes de otimização; modificando memória, processador, entre outros, até alcançar um cenário que seja satisfatório de acordo com sua necessidade.

A organização que utilizará o VMware Capacity Planner precisa disponibilizar um servidor com o sistema operacional Windows Server 2003 Standard ou superior, uma lista com os servidores alvos da VIRTUALIZAÇÃO e portas apropriadas, acesso ao domínio e uma conexão VPN. O VMware Capacity Planner Tool permite simplificar o processo de avaliação e fornecer dados que permitem definir o cenário ideal para a consolidação. O VMware Capacity Planner Tool também permite modelar cenários de consolidação.

A Microsoft também possui uma ferramenta equivalente de Planejamento da Capacidade intitulada Microsoft Assessment Planing Toolkit, que pode ser encontrada em http://technet.microsoft.com/en-us/library/bb977556.aspx

Prova de Conceito (*Proof of Concept – POC*)

60 servidores a 360 servidores: neste caso, o mais prudente é desenvolver uma prova de conceito (POC – *proof of concept*) do ambiente a ser virtualizado junto com análise de TCO. Utilizar também uma ferramenta do tipo VMware Capacity Planner Tool. A solução deve ser discutida e refinada com o cliente. As recomendações devem ser feitas através de um *Virtualization Readiness Assessment* (VRA), que define os servidores a serem incluídos na VIRTUALIZAÇÃO e os servidores que devem ser excluídos, possíveis cenários de consolidação e considerações sobre o uso da SAN.

Uma Prova de Conceito é desenhada para prover às organizações um conhecimento das soluções de infraestrutura virtual e demonstrar as capacidades da tecnologia existente nos software e como são aplicadas para resolver problemas reais.

A prova de conceito, geralmente, demora 30 dias e necessita o envolvimento de um engenheiro de sistemas da fornecedora de software ou de um parceiro, auxiliando a organização:

- Na instalação e configuração do software.
- Definição das melhores práticas para o desenho da solução, implementação e utilização.
- Teste de características e funcionalidades do software e como este se aplica a diferentes cenários de utilização.
- Validar que a solução proposta vai ao encontro das necessidades do negócio da organização.

POC Impact Report

O *POC Impact Report* é um documento elaborado pelo consultor técnico do parceiro / Engenheiro de Sistemas da fornecedora de software que inclui informações técnicas e referentes ao negócio do cliente e também sobre o impacto da solução no ambiente do cliente.

O projeto, como qualquer outro, deve utilizar uma ferramenta básica de software que permita gerenciar os prazos e recursos para a realização das duas principais entregas definidas no POC e descritas abaixo:

- **Especificações de sistema**: devem estar incluídas as especificações detalhadas, os requisitos de licenciamento, requisitos de backup, implicações da VIRTUALIZAÇÃO na rede SAN.
- **Plano de Implementação**: o plano de implementação deve incluir a verificação dos recursos, construção da infraestrutura, migrações, revisão.
- Deve haver uma estratégia clara de comunicação.

Também o funcionamento da nova infraestrutura deve ser cuidadosamente planejado, envolvendo testes das principais aplicações, antes de estarem em produção, que estarão funcionando no ambiente virtualizado. Por mais que os serviços de implementação tenham sido contratados, deve-se pensar no treinamento de pelo menos um profissional interno para mitigação dos riscos com a nova operação. Esta opção deve fazer parte de um plano maior de estratégia para redução de riscos.

9.8.11. Conceito na Prática: Microsoft SQL Server

Segundo a Microsoft, os recursos do Microsoft SQL Server 2008 suprem três áreas principais:

- **Flexibilidade**: a solução de consolidação do SQL Server 2008 pode consistir de diversos bancos de dados em uma única instância de Servidor SQL, em somente um servidor físico, várias instâncias do SQL Server em um computador físico ou diversos servidores virtuais em um computador físico. Com o suporte a todos esses métodos de consolidação, o SQL Server 2008 possibilita às organizações

escolher o grau de isolamento apropriado para o desempenho necessário. O SQL Server 2008 também fornece diversas ferramentas de migração dos dados e bancos de dados existentes para um servidor consolidado.

- **Gerenciamento**: embora o propósito principal da consolidação de servidor seja a redução dos custos de hardware e licenciamento, esta também pode ser utilizada na centralização de funções administrativas. O SQL Server 2008 fornece um conjunto de ferramentas para gerenciar, administrar, monitorar e solucionar problemas dos sistemas de dados consolidados. Essas ferramentas habilitam a administração centralizada dos servidores consolidados, além de permitir a consolidação da função de gerenciamento para servidores distintos.
- **Escalabilidade e desempenho**: ao consolidar os sistemas de dados em menos servidores, cada um deles se encarregará do aumento da carga de trabalho. Hardwares de melhor desempenho talvez solucionem alguns problemas, todavia outros podem ocorrer no momento em que alguns dos processos do servidor de consolidação utilizam uma quantidade de recursos desproporcional, ocasionando um desempenho menor do que esperado nos outros processos. Os problemas também ocorrem quando um processo bloqueia os recursos de outro processo, impedindo a sua conclusão. O SQL Server 2008 inclui os recursos de otimização do desempenho, os quais podem auxiliar na solução desses problemas.

Consolidação das Instâncias do SQL Server

A abordagem mais simples para consolidar os serviços de dados com o MS SQL Server 2008 consiste no uso de uma única instância do SQL Server com vários bancos de dados. Esta abordagem é apropriada quando todos os bancos de dados possuem requisitos de segurança, gerenciamento e compatibilidade semelhantes. Além disso, o hardware deve fornecer o nível de desempenho e escalabilidade necessário às cargas de trabalho geradas em todos os bancos de dados. Esta situação se torna cada vez mais comum. O MS SQL Server 2008 pode utilizar o Windows Server 2008 ou o Windows Server 2003.

O número máximo de instâncias suportadas por cada edição do SQL Server 2008 é apresentado na **Tabela 9-10** a seguir:

Tabela 9-10 Instâncias do SQL Server 2008

Edição	Número máximo de instâncias
SQL Server 2008 Standard Edition	16
SQL Server 2008 Enterprise Edition	50
SQL Server 2008 Developer Edition	50

Quando os bancos de dados possuem diferentes requisitos de segurança, gerenciamento e compatibilidade, os serviços de dados podem ser consolidados através da execução de diversas instâncias do SQL Server 2008 simultaneamente em um único servidor físico, com o intuito de reduzir os custos de hardware e licenciamento, bem como as despesas administrativas. As instâncias estão completamente isoladas umas das outras, de forma que alterações em uma delas não afetam as outras instâncias localizadas no mesmo computador. Além da redução dos custos de hardware pela consolidação, obtêm-se benefícios com a diminuição dos custos de licenciamento, pois só será necessária uma licença para cada processador físico, independentemente de quantas instâncias estão instaladas.

A fim de obter um isolamento completo ao nível do sistema operacional, o SQL Server 2008 suporta também a VIRTUALIZAÇÃO de servidor. Ao consolidar os serviços de dados com a VIRTUALIZAÇÃO, obtêm-se níveis de isolamento adequados entre soluções de bancos de dados com diferentes WORKLOADS, requisitos de segurança, gerenciamento e compatibilidade, reduzindo ao máximo o número de servidores e licenças necessários, simplificando a infraestrutura da rede.

A VIRTUALIZAÇÃO com vSphere possibilita obter taxas de consolidação com o MS SQL Server da ordem de 4x até 20x, segundo a VMware. Cada VM pode escalar até 8vCPUS e 255GB de RAM e suportar cargas intensivas em I/O. A VIRTUALIZAÇÃO permite otimizar o uso dos recursos multicore dos processadores distribuindo a carga de múltiplos bancos de dados.

Os bancos de dados são provisionados para cima por questões de segurança, mas boa parte deles utiliza poucos recursos do servidor. Qual a razão para não serem virtualizados? A VIRTUALIZAÇÃO pode reduzir os custos de hardware, software e de licenças. As funcionalidades do vSphere como HA, DRS e FT podem ser utilizadas com o MS SQL Server. O licenciamento da Microsoft agora atende à demanda do vMotion e facilita o seu uso. A exceção da VIRTUALIZAÇÃO é quando se tem um ambiente de produção com o SQL SERVER utilizando WORKLOADS que consomem todo o recurso do servidor. Em ambientes de teste e desenvolvimento a VIRTUALIZAÇÃO, pode-se assim dizer, é obrigatória.

A Microsoft suporta o MS SQL Server, Exchange e SharePoint Server em ambientes vSphere. O VMware ESX foi o primeiro HYPERVISOR validado pelo Microsoft Virtualization Validation Program (SVVP).

A **Figura 9-17** ilustra as três possibilidades de consolidação. A opção (a) trata da consolidação em um única instância, a opção (b) representa a consolidação de um único banco de dados de várias instancias e a opção (c) ilustra a consolidação com a VIRTUALIZAÇÃO.

A consolidação de servidor geralmente requer o movimento de bancos de dados, aplicativos e sistemas operacionais, de fontes existentes para servidores consolidados. O SQL Server 2008 fornece ferramentas e tecnologias que podem ser utilizadas para migrar dados e bancos de dados, o que compreende fazer o backup e restaurar, bem como utilizar o *SQL Server Integration Services*. Para auxiliar na migração de bancos de dados

antigos para uma plataforma SQL Server 2008 consolidada, existe também o suporte à atualização direta dos bancos de dados do SQL Server 2000 e SQL Server 2005.

Os discos físicos no *storage* propiciam capacidade de *storage* ao MS SQL Server. Porém, antes dos dados serem carregados, os discos precisam ser configurados. Os grupos de disco precisam de níveis de RAID adequados. Também a criação das LUNs deve obedecer ao requisito de otimizar a performance.

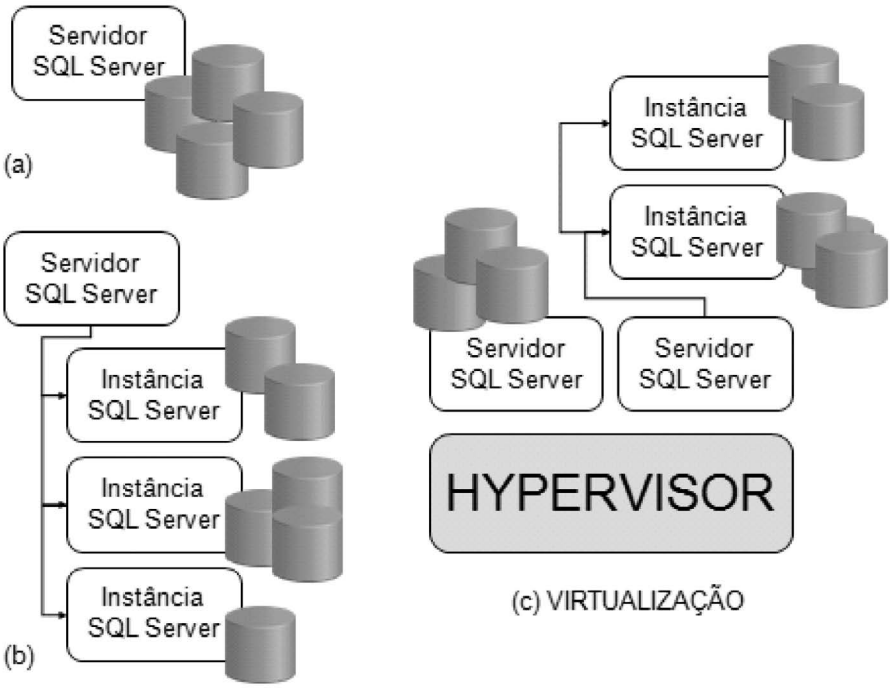


Figura 9-17 Consolidação do MS SQL Server 2008

Os componentes do MS SQL Server são:

- **Tempdb**: é uma área de trabalho compartilhada por todos os bancos de dados no servidor e para atividades como criar tabelas temporárias e queries.
- ***Transaction log:*** armazena os detalhes das modificações feitas no banco de dados com todos os detalhes das transações que realizam as modificações.
- **Dados**: objetos de dados definidos pelo usuário, índices e storage procedures.

A Microsoft recomenda algumas práticas para o MS SQL Server:

- Colocar tempdb e *transaction logs* em discos físicos separados. Significa separar estes arquivos em seus próprios discos virtuais em disk groups separados.
- Espalhar os objetos do banco de dados por vários splindles.
- Utilizar RAID 10 como configuração de RAID ideal para a manutenção dos dados do MS SQL Server, pois este nível de RAID propicia alta disponibilidade e alto desempenho. Na prática, por questões de custo, pode-se utilizar RAID 0 para o Tempdb, RAID1 para os logs de transação e RAID 5 para os dados.

Otimização do Desempenho com o SQL Server 2008

Devido ao crescente aumento dos dados corporativos em tamanho e complexidade, é necessário tomar providências de forma a otimizar o tempo de acesso aos dados.

O SQL Server 2008 beneficia-se dos sistemas de 64 bits, multicore e multiprocessadores. Para suportar a ampliação de cargas de acesso a dados, relatórios e cargas de conteúdo analítico, o SQL Server pode operar com uma memória de até 64 GB e suportar a alocação dinâmica de memória mapeada pelo AWE (Address Windowing Extensions) em um hardware de 32 bits, além de utilizar uma memória de até 8 terabytes em um hardware de 64 bits.

O SQL Server 2008 inclui diversos recursos e aprimoramentos para a otimização do desempenho em todas as áreas de funcionalidade, incluindo o seu uso como banco de dados relacional OLTP (*Online Transaction Processing*), banco de dados OLAP (*Online Analytical Processing* – Processamento Analítico Online) e na execução de relatório e processos de dados ETL (*extraction, transformation, and loading* – extração, transformação e carga).

A consolidação do servidor, as armazenagens volumosas de dados e as consultas complexas exigem recursos físicos, a fim de suportar as diversas WORKLOADS que estão sendo executadas no servidor. O SQL Server 2008 possui a habilidade de se beneficiar das mais recentes tecnologias de hardware. Diversas instâncias de mecanismos de banco de dados e instâncias do *analysis services* podem ser instaladas em um único servidor para consolidar o uso do hardware. Segundo a Microsoft, até 50 instâncias podem ser instaladas em um único servidor sem comprometer o desempenho ou a capacidade de resposta.

O uso do MS SQL Server se dá da seguinte forma:

- **OLTP**: os bancos de dados relacionais são parte central dos serviços e aplicativos críticos na maioria dos ambientes empresariais. As empresas devem ser capazes de assegurar a consistência do desempenho e da capacidade de resposta de seus sistemas de dados à medida que o volume de dados cresce, bem como o número de usuários e aplicativos que dependem do armazenamento de dados relacionais. Segundo a Microsoft, o SQL Server 2008 fornece um mecanismo de banco de dados que suporta bancos de dados relacionais volumosos e complexos processamentos de consulta. O mecanismo de processamento de consulta do SQL Server contribui para o usuário maximizar o desempenho dos aplicativos. Este mecanismo avalia solicitações de modo a gerar planos de execução de consultas ideais, baseados em estatísticas sobre os índices, seletividade padrão e volumes de dados, sustentados de forma dinâmica. No SQL Server 2008 é possível bloquear esses planos de consultas para assegurar um desempenho consistente com as consultas frequentemente executadas. O mecanismo de processamento de consultas também pode tirar proveito dos servidores com diversos processadores e núcleos de forma a gerar planos de execução que se beneficiam do paralelismo, a fim de aumentar ainda mais o desempenho.

Em geral, a operação mais impactante em termos de desempenho de consulta é a I/O de disco. Os recursos de cache dinâmico do SQL Server reduzem a quantidade necessária de acesso físico ao disco para recuperação e modificação dos dados. Assim, o mecanismo de processamento de consulta pode aprimorar o desempenho através do uso de varreduras "read-ahead", com o intuito de antecipar as páginas de dados necessárias a um plano de execução determinado e fazer a leitura de forma preemptiva no próprio cache. Além disso, o suporte nativo à compressão de dados do SQL Server 2008 pode reduzir o número de páginas de dados que precisam ser lidas, o que melhora o desempenho das cargas de trabalho associadas a I/O.

O SQL Server 2008 suporta o particionamento de tabelas e índices que permite aos administradores controlar a disposição física dos dados através da atribuição de partições a partir da mesma tabela ou índice, para vários grupos de arquivos em dispositivos de armazenagem física distintos. Com as otimizações feitas ao mecanismo de processamento de consulta do SQL Server 2008, é possível colocar o acesso aos dados particionados em paralelo, o que aperfeiçoa o desempenho.

No SQL Server 2008 estão incluídos o SQL *Server Profiler* e o *Database Engine Tuning Advisor*. Com o uso do SQL Server Profiler é possível capturar o rastreio dos eventos que ocorrem em uma carga de trabalho típica para o aplicativo e, em seguida, repetir o rastreamento no *Database Engine Tuning Advisor*, que gera e implementa recomendações de indexação e particionamento de dados, de forma a possibilitar a otimização do desempenho do aplicativo.

Após criar os índices e as partições que melhor se adaptem à carga de trabalho do aplicativo, é possível utilizar o SQL Server Agent para programar um plano de manutenção de banco de dados automático. A manutenção automática reorganiza e reconstrói os índices periodicamente, além de atualizar as estatísticas de índice e de seletividade, a fim de assegurar o desempenho otimizado de maneira consistente, à medida que as inserções e modificações de dados fragmentam as páginas de dados físicas do banco de dados.

Muitas vezes utiliza-se um único servidor, a fim de fornecer vários serviços de dados. Em alguns casos, diversos aplicativos e cargas de trabalham recorrem à mesma fonte de dados. À medida que continua a tendência atual de consolidação de servidores, pode ser difícil fornecer um desempenho previsível para uma carga de trabalho determinada, pois outras cargas de trabalho no mesmo servidor competem pelos recursos do sistema. Devido às diversas cargas de trabalho em um único servidor, os administradores devem evitar problemas como, por exemplo: a consulta de fuga que se alimenta de outra carga de trabalho de recursos do sistema ou as cargas de trabalho de baixa prioridade que afetam de maneira adversa as cargas de trabalho de alta prioridade. O SQL Server 2008 inclui o Resource Governor, que permite aos administradores definir limites e atribuir prioridades a cargas de trabalho individuais sendo executadas em uma instância do SQL Server. As cargas de trabalho baseiam-se em fatores como usuários, aplicativos e bancos de dados.

Através da definição de limites aos recursos, os administradores podem minimizar a possibilidade de ocorrência de consultas de fuga, bem como restringir os recursos

disponíveis para as cargas de trabalho que os monopolizam. Ao determinar prioridades, é possível otimizar o desempenho dos processos de missão crítica e ao mesmo tempo manter a previsibilidade das outras cargas de trabalho do servidor.

O SQL Server 2008 fornece também o Performance Studio, uma estrutura integrada que pode ser utilizada em coleta, análise, solução de problemas e armazenagem de informações de diagnóstico do SQL Server. O Performance Studio oferece a solução de monitoramento de desempenho.

- **OLAP e DATA WAREHOUSE (DW):** os ambientes de data warehouse devem se adaptar ao crescente volume de dados e requisitos de usuários, mas manter o ótimo desempenho. Conforme as consultas de data warehouse tornam-se mais complexas, cada parte da consulta deve ser otimizada para manter um desempenho aceitável. No SQL Server 2008, o otimizador de consulta pode introduzir de maneira dinâmica um filtro bitmap otimizado, a fim de aprimorar o desempenho da consulta para as consultas de esquema star. Além disso, o SQL Server 2008 suporta o particionamento de dados, a funcionalidade de indexação avançada e modos de exibição indexados a fim de sustentar armazenamentos de dados mais volumosos, colunas esparsas e tipos de dados eficientes. Os DATA WAREHOUSES são tipicamente utilizados por diversos consumidores de dados somente leitura, do tipo soluções de análise e relatório. Estes podem ficar sobrecarregados com solicitações de dados, o que reduz a capacidade de resposta. Para superar esse problema, o SQL Server 2008 suporta bancos de dados compartilhados e escalonáveis, fornecendo uma maneira de fazer a escalabilidade horizontal dos bancos de dados de relatório somente leitura em múltiplas instâncias de servidor de bancos de dados, a fim de distribuir a carga de trabalho do mecanismo de consulta e isolar as consultas de muitos recursos. O recurso de banco de dados compartilhado e escalonável permite aos administradores criar uma fonte de dados somente de leitura dedicada, através da montagem de cópias de banco de dados somente leitura em diversos servidores de relatórios. Os aplicativos acessam uma cópia consistente dos dados, independentemente do servidor de relatório ao qual os aplicativos se conectam.

Os aplicativos de *Analysis Services* geralmente requerem cálculos longos e complexos. Um tempo precioso de processamento é desperdiçado com o cálculo de agregações cuja resolução é NULL ou zero. Os cálculos bloqueados do SQL Server 2008 Analysis Services usam valores padrão, minimizam o número de expressões que precisam ser computadas e limitam a navegação da célula a uma vez para o espaço inteiro, em vez de uma vez para cada célula, o que melhora de maneira significativa o desempenho de cálculo.

Embora as partições do MOLAP (OLAP Multidimensional) forneçam um melhor desempenho de consulta, anteriormente as empresas que necessitavam dos recursos de write-back precisavam utilizar as partições do ROLAP (OLAP Relacional) para manter as tabelas de write-back. O SQL Server 2008 adiciona a capacidade de executar operações de write-back nas partições MOLAP, eliminando a degradação de desempenho, devido à manutenção das tabelas de write-back do ROLAP.

O mecanismo Reporting Services do SQL Server 2008 adiciona desempenho e escalabilidade ao Reporting Services, por meio do processamento sob demanda. Os relatórios não são mais ligados à memória porque o processamento de relatório agora utiliza o cache do sistema de arquivos para se adaptar à pressão desta. O Report Processing também pode se adaptar a outros processos que consomem memória. Uma nova arquitetura de renderização remove os problemas de uso de memória de versões de renderizadores anteriores.

- **ETL:** os processos ETL são frequentemente utilizados para popular e atualizar os dados corporativos dos bancos de dados de origem dos data warehouses por toda a empresa. Tradicionalmente, muitas empresas necessitam somente de dados históricos com renovações de dados infrequentes do data warehouse. Atualmente, muitas organizações necessitam de dados quase que em tempo real disponibilizados pelo datawarehouse. À medida que aumenta a quantidade de dados e as renovações de dados no data warehouse precisam ser mais frequentes, o tempo de processo e a flexibilidade do ETL tornam-se mais importantes. As renovações de dados exigem que o SQL Server Integration Services utilize pesquisas para comparar as linhas de origem aos dados no interior do data warehouse. O *Integration Services* inclui o desempenho de pesquisa, que diminui o tempo de execução de pacotes e otimiza as operações ETL. Outro problema com os processos ETL tradicionais consistia em determinar quais dados sofreram alterações no banco de dados de origem. Os administradores precisavam ser extremamente cuidadosos para evitar a duplicação dos dados existentes. Alguns administradores preferem remover todos os valores de dados e recarregar o data warehouse em vez de gerenciar os dados modificados. Isso aumentava a sobrecarga ao processo ETL. O SQL Server 2008 inclui a funcionalidade do CDC (*change data capture*) para registrar as atualizações de alteração de tabela, que auxilia no acompanhamento de alterações dos dados, bem como garante a consistência do data warehouse quando as renovações de dados são programadas.

Sizing do MS SQL Server

Para a realização do sizing do MS SQL Server 2008 é necessário saber inicialmente se o banco de dados será utilizado para OLTP ou para sistemas de data warehouse.

Se o *sizing* é para uma solução OLTP, precisam ser definidos:

- O tipo de processador a ser utilizado.
- O tipo de storage a ser utilizado (SAS ou FC).
- O tamanho da base de dados (atual e para os próximos 3 anos).
- As funcionalidades a serem contempladas.
- O máximo número de usuários concorrentes/camada de aplicação.
- O máximo de transações por minuto.
- O tamanho do Tempdb.
- A utilização do processador em %.

Se o *sizing* é para uma solução de data warehouse, precisam ser definidos:

- O tipo de processador a ser utilizado.
- O tipo de storage a ser utilizado (SAS ou FC).
- O tamanho da base de dados (atual e para o próximos 3 anos).
- As funcionalidades a serem contempladas.
- A quantidade média das linhas das tabelas.
- A quantidade média de arquivos afetados por uma query.
- A quantidade média de queries submetidas por hora.
- O tamanho do Tempdb.
- A utilização do processador em %.

Nos dois casos é necessário considerar o overhead introduzido pelo HYPERVISOR. Por segurança, deve-se considerar 15% de overhead, ou seja, o servidor físico definido pela ferramenta de sizing deve ser 15% maior se o ambiente for virtualizado.

9.8.12. Conceito na Prática: Oracle Database

Oracle Database

O Oracle Database 11G (G é de grid) significa:

- Beneficiar-se dos componentes disponíveis para o grid.
- Balanceamento de carga.
- Compartilhar informações independentes da localização.
- Agendar recursos no grid.

As versões 8, 9 e 10 podem ser atualizadas diretamente para o 11G. A Oracle fornece o Oracle Database Upgrade Assistant (DBUA) como ferramenta de atualização.

Oracle ASM

O Oracle Database 10G introduziu também o *Automatic Storage Management,* um serviço que fornece gerenciamento das unidades de disco. O ASM é uma alternativa ao uso de arquivos brutos não formatados (*raw*) ou processados (*cooked*).

O ASM permite:

- Compatibilidade com configurações de disco tipo JBOD ou SAN.
- Prevenção contra exclusão acidental de arquivos.
- Balanceamento de cargas pelas unidades de disco.
- Capacidade de espelhar dados em diferentes discos.

Oracle RAC

O Oracle 9i Database introduziu o *Real Application Cluster*s (RAC), sucessor do Oracle Parallel Server. O Oracle Database 11G oferece (relativo ao RAC):

- **Clusterware portável**: elimina a necessidade de um clusterware oferecido pelo fornecedor.
- **Atualizações sem interrupção**: permite atualizar patches sem interrupção, com algumas exceções.
- **Cluster Ready Services**: fornece serviços adicionais de gerenciamento para o cluster como associação de nós, gerenciamento global e alta disponibilidade.
- Paralelismo de recuperação aprimorado em sistemas com múltiplas CPUs.

O Oracle RAC é uma solução de banco de dados em *cluster* que requer dois ou mais nós de hardware capazes de trabalhar juntos sob o controle de um sistema operacional. A arquitetura utilizada é do tipo shared cache. O RAC é a única solução ativa-ativa baseada em plataforma x86 existente e separa as instâncias do banco de dados em si, o que não é tão comum de ser notado com a solução de *cluster* de uma instância/banco de dados únicos. O Oracle RAC permite alto poder de processamento e alta disponibilidade em um único *cluster*. Na versão 11G, dependendo da versão do Oracle Database e do hardware a ser utilizado, o Oracle RAC pode ser até gratuito, ou seja, viria como uma funcionalidade do Oracle Database.

Em uma arquitetura de n-camadas, comum nos DATACENTERS com foco empresarial, o Oracle RAC seria a opção natural para propiciar alto desempenho na camada de banco de dados, pois o processamento de cada nó é somado ao processamento final, com baixíssima perda de performance. Em certas instalações de ERP, por exemplo, esta seria a saída natural para manter as n camadas em baixa plataforma.

O Oracle Cluster File System (OCFS2) facilita a administração do Oracle RAC, introduzindo uma imagem consistente do sistema de arquivos através dos servidores em um cluster.

Sizing Oracle

Para a realização do sizing do Oracle 10G & 11G, é necessário saber inicialmente se o banco de dados será utilizado para uma função de OLTP ou para uma função de apoio a decisão baseados em *data warehouse*.

O sizing, se realizado para OLTP, deve contemplar:

- Tipo de processador a ser utilizado (AMD, Intel).
- Tipo de servidor a ser utilizado (RACK, blade).
- Tipo de storage a ser utilizado (SAS , iSCSI ou FC).
- Tamanho da base de dados (agora e próximos 3 anos).
- Funcionalidades a serem contempladas.

- Máximo número de usuários concorrentes/camada de aplicação.
- Máximo de transações por minuto.
- Utilização do processador em %.

O sizing, se realizado para suporte a decisão (SSD) e data warehouse, deve contemplar as seguintes definições:

- Tipo de processador a ser utilizado (AMD, Intel).
- Tipo de servidor a ser utilizado (torre, RACK, blade).
- Tipo de storage a ser utilizado (SAS, iSCSI ou FC).
- Tamanho da base de dados (agora e próximos 3 anos).
- Funcionalidades a serem contempladas.
- Throughput de I/O.
- Utilização do processador em %.

É necessário considerar o overhead introduzido pelo HYPERVISOR. Por segurança deve-se considerar 15% de overhead, ou seja, o servidor físico definido pela ferramenta de sizing deve ser 15% maior se o ambiente for virtualizado.

A VIRTUALIZAÇÃO do Oracle DB (single instance/RAC) é uma questão relevante. A opção de virtualizar o Oracle RAC com o vSphere, por exemplo, não é suportada pela Oracle. Já a VIRTUALIZAÇÃO de uma instância única é suportada e existem várias razões para fazê-la. Senão, vejamos: a VIRTUALIZAÇÃO permite rodar vários múltiplos bancos de dados Oracle em um único servidor físico. Também permite que vários ambientes de teste Oracle compartilhem o mesmo servidor físico e permite consolidar vários ambientes de bancos de dados em poucos servidores físicos.

O paper *VMware and Oracle Databases Software Solutions Deployment Guide,* publicado pela VMware e disponível na Internent, dá uma série de dicas de como pensar a VIRTUALIZAÇÃO do ambiente Oracle com o VMware. O vMotion, por exemplo, pode facilitar a migração de um banco de dados Oracle de um servidor físico para outro.

9.8.13. Conceito na Prática: Microsoft Exchange Server

Introdução

O MS Exchange Server 2010/2007 é uma plataforma de correio de 64 bits e roda em hardware x86 64 bits. Pode endereçar acima de 4GB de memória e ter até 50 *storage groups*. Também os recursos de cluster estão mais sofisticados nesta versão. As novas características do vSphere permitem rodar o Exchange Server para grandes instâncias de mailbox.

- As VMs podem escalar até 8 vCPU e 255 GB de memória.
- A escalabilidade de I/O do VMware vSphere Disk IO aumentou para mais de 350.000 IOPS, possibilitando que o VMware ESX suporte aplicações intensivas em I/O como o Exchange.
- O VMware vSphere suporta I/O de rede de até 40 Gbps.

Um bom exemplo de ganho de escalabilidade com o Exchange em ambiente virtualizado é baseado no teste que a VMware fez. Sem o VMware, um servidor de mailbox Exchange 2007 rodando em um servidor físico conseguiu ir bem até 8000 *heavy users* de mailbox. Não adianta utilizar maiores servidores, pois isto não resolve o problema da escalabilidade. Com o VMware, o Exchange pode ser escalado em 8 máquinas virtuais, cada uma com 2.000 mailbox de *heavy users* e chegar a 16.000 usuários em um servidor com 16 cores.

Visão Geral de Funções de Servidor Exchange

Uma função de servidor é uma unidade que agrupa logicamente os recursos e componentes necessários para executar uma função específica no ambiente de mensagem. O requisito de uma função de servidor é um servidor que poderá ser executado como uma unidade atômica de escalabilidade. Uma função de servidor é composta por um grupo de recursos.

As funções de servidor permitem que os administradores escolham facilmente quais recursos são instalados em um servidor Exchange. O agrupamento lógico de recursos em funções de servidor oferece as seguintes vantagens:

- Reduz a superfície de ataque em um servidor Exchange.
- Permite instalar e configurar um servidor Exchange da maneira que se pretende usá-lo.
- Oferece uma instalação simples e a capacidade de personalizar completamente um servidor para oferecer suporte a metas e necessidades comerciais.

O Exchange Server 2007 tem as seguintes funções de servidor:

- **Servidor Mailbox (Caixa de Correio):** é um servidor backend que pode hospedar caixas de correio e pastas públicas.
- **Servidor Client Access (Acesso para Cliente):** é o servidor de camada intermediária que hospeda os protolocos de cliente, como Post Office Protocol 3 (POP3), Internet Message Access Protocol 4 (IMAP4), Secure Hypertext Transfer Protocol (HTTPS), Outlook em Qualquer Lugar, serviço de Disponibilidade e serviço Autodiscover. O Servidor de Acesso para Cliente também hospeda serviços da Web.
- **Servidor Unified Messaging (Unificação de Mensagens):** é o servidor de camada intermediária que conecta um sistema PBX (Private Branch eXchange) ao Exchange 2007.
- **Servidor HUB Transport (Transporte de Hub):** é o servidor de roteamento de email dentro da organização do Exchange.

- **Servidor de Edge Transport (Transporte de Borda):** É o servidor de roteamento de email que geralmente está localizado no perímetro da topologia e roteia emails internos e externos da organização do Exchange.

Estas funções se integram ao Microsoft AD com exceção do Edge Transport Server, que, inclusive, não pode ser consolidado mesmo em pequenas instalações. Quando o servidor de Mailbox for configurado em cluster, nenhuma outra função pode ser configurada nestes servidores.

Topologia Física

Pode-se classificar a organização do Microsoft Exchange de três maneiras:

- **Topologia lógica:** refere-se à estrutura da floresta do serviço de diretório do Active Directory.
- **Topologia física:** refere-se aos locais físicos e aos computadores da organização do Exchange.
- **Topologia da organização do Exchange:** refere-se às funções de servidor do Exchange e ao modo de posicionamento dentro das topologias física e lógica. Interessa para o contexto do livro a topologia física.

A topologia física para o MS Exchange Server 2007 mapeia elementos físicos para locais geográficos. Uma topologia física geralmente é usada para descrever uma rede ou local de servidores. Geralmente, todas as topologias físicas se baseiam em requisitos organizacionais específicos e únicos para recursos de escopo baseados nos requisitos comerciais e de segurança.

A topologia física também classifica a distribuição de servidores e as funções de gerenciamento em duas categorias principais: servidores e administração centralizados, e servidores e administração distribuídos.

Se a organização for composta por sites remotos que estejam todos conectados por banda larga e conexões de rede confiáveis, independentemente da distância entre os escritórios, pode-se implantar um sistema de mensagens centralizado. Um sistema de mensagens centralizado significa que todos os servidores que estão executando o Exchange estão localizados e são gerenciados em um único DATACENTER. Ao planejar o sistema de mensagens, é melhor começar considerando este modelo, pois ele é o mais econômico e fácil de gerenciar.

Se a organização possui sites remotos com pouca largura de banda, alta latência ou conexões de rede não confiáveis, pode-se introduzir servidores para controlar o roteamento do tráfego de mensagens de um local para outro. No entanto, os locais remotos e os múltiplos grupos de roteamento não o impedem de centralizar o modelo administrativo. Além disso, com os recursos do Microsoft Windows Server 2008, do Exchange 2007 e do Microsoft Office Outlook 2007, pode-se também consolidar o hardware dos servidores removendo servidores que estejam executando o Exchange a partir de sites remotos. Com essas alterações, os usuários podem fazer logon remotamente nos servi-

ços do Microsoft Windows e no Exchange 2007 e ter menos problemas relacionados a uma redução no desempenho ou na conectividade.

Características de um Sistema de Mensagens Centralizado

Um sistema de mensagens centralizado consiste em um grande DATACENTER que hospeda todos os recursos do servidor, incluindo o serviço de diretório *Active Directory*, os servidores de catálogo global, os controladores de domínio e os servidores Exchange. O DATACENTER dá suporte a todos os usuários do sistema de mensagens, quer estejam conectados localmente ou remotamente. A **Figura 9-18** ilustra a arquitetura de um sistema de mensagens centralizado baseado no MS Exchange Server 2007. Observa-se as opções de conexão da rede PBX/VoIP. Dependendo do *sizing*, esta arquitetura varia para cima e para baixo. Com a VIRTUALIZAÇÃO, optou-se por ter só dois servidores de Mailbox. A rede de armazenamento pode ser baseada em FC ou iSCSI.

As características a seguir referem-se a um sistema de mensagens centralizado:

- Os dados são hospedados e gerenciados em um local centralizado, independentemente de os usuários estarem conectados remotamente. Isso contrasta com o modelo distribuído, em que os usuários têm acesso local a caixas de correio, mas a administração do servidor é mais complexa.
- As atualizações de software podem ser feitas a partir de um local centralizado.

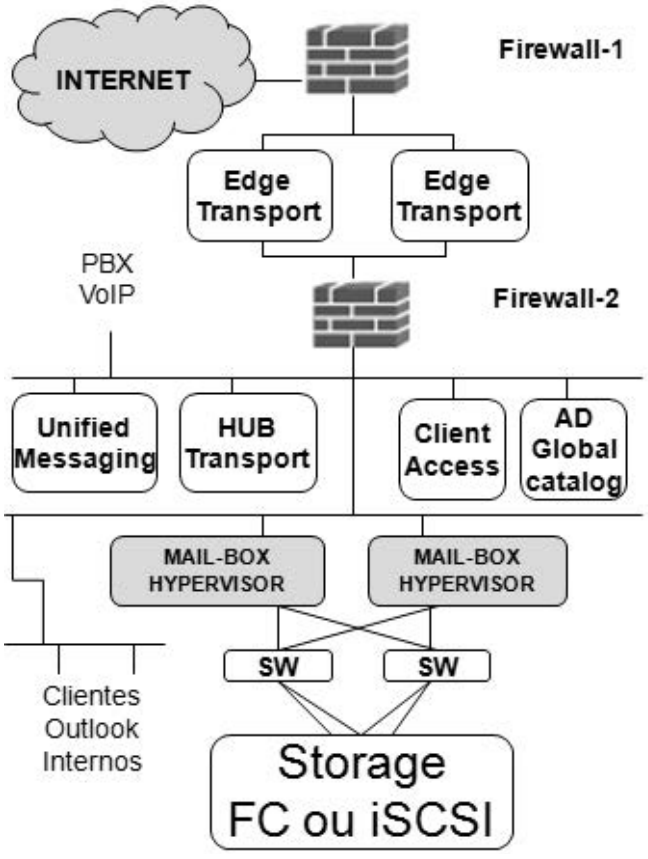


Figura 9-18 Arquitetura para MS Exchange

Os requisitos comerciais associados à redução de requisitos de custo e de segurança são a força motriz por trás dos sistemas centralizados. Os requisitos giram pela centralização local (reduzindo o número de sites que fornecem recursos do servidor), consolidação física (substituindo servidores menores por servidores de última geração), consolidação administrativa e consolidação de dados (centralizando soluções de armazenamento que fornecem capacidades de recuperação de desastres e backup).

O *design* centralizado estabelece que os pré-requisitos das seguintes áreas sejam atendidos ou estejam incluídos no plano do projeto:

- **Custos de hardware do DATACENTER**: compare o custo de instalação de servidores e clusters de última geração do DATACENTER com a economia do custo administrativo da centralização de servidores. Recomenda-se agrupar os servidores backend para criar alta disponibilidade e redundância no sistema, mas esta opção envolve custos iniciais mais altos. No entanto, esses custos podem ser compensados pelas reduções nos custos operacionais, custos de infraestrutura, redução da inatividade e maior escalabilidade.
- **Planejamento de contingência**: ao centralizar recursos do servidor e de dados na organização, aumenta-se o número de possíveis pontos únicos de falha. Você deve fazer planos de contingência, no caso de um evento catastrófico afetar o seu DATACENTER.
- **Interrupções na rede**: considere o impacto que uma interrupção na rede terá sobre os usuários em locais remotos. Se os usuários tiverem o Modo do Exchange em Cache habilitado no Outlook, essa consideração será um problema pequeno.
- **Reduções nos custos operacionais e administrativos**: a centralização dos recursos do servidor pode reduzir custos operacionais, pois a capacidade de serviço e o crescimento são obtidos em virtude de um melhor uso dos recursos. Ela reduz também os custos de infraestrutura associados aos requisitos de armazenamento e backup.
- **Armazenamento de dados**: com maiores volumes de dados centralizados, deve-se utilizar sistemas de armazenamento mais confiáveis para melhorar a integridade dos dados. Além disso, reduzindo a complexidade da infraestrutura do servidor, é possível restaurar serviços e dados com mais facilidade, no caso de ocorrerem falhas.
- **Conectividade de LAN e WAN**: se sua rede atual não fornece a largura de banda e a velocidade necessárias para centralizar servidores, será necessário fazer uma atualização na rede no plano de projeto.
- **Segurança**: um modelo centralizado facilita o gerenciamento da segurança e, portanto, oferece mais controle. Esse controle proporciona à equipe de segurança mais facilidade para manter assinaturas de verificador de vírus atualizadas e executar ações adequadas em resposta a incidentes de seguranças. Outra vantagem de um design centralizado é que ele localiza os seus servidores em um DATACENTER que pode ser protegido fisicamente.

Sizing do MS Exchange

O *sizing* do MS Exchange 2007 é complexo e devem ser consideradas as opções de funções do servidor a serem implementadas.

Inicialmente, deve-se definir que aspectos devem ser endereçados e se a otimização deve ser para preço ou para desempenho.

Segundo a VMware, o vSphere permite atingir taxas de consolidação da ordem de 5x ou 10x com o Exchange Server sem precisar utilizar a complexidade do Micrososft clustering. A migração para hosts maiores com o MS Exchange pode ser feita facilmente com o VMware VMotion. Também novos serviços como o BlackBerry Enterprise Server podem ser provisionados em pouco tempo. O uso do vMotion permite obter a alta disponibilidade e a recuperação de desastres sem servidor exclusivo stand-by.

Importantes definições para uso de ferramentas de sizing de qualquer fabricante:

- Tipo de processador a ser utilizado (AMD, Intel).
- Tipo de servidor a ser utilizado (RACK, blade).
- Tipo de storage a ser utilizado (SAS, iSCSI ou FC).
- Os números de caixas postais devem ser estimados para três tipos de usuários (tiers), tanto em termos de tamanho da caixa postal como do profile de I/O.
- Definir o crescimento do número de caixas postais para os próximos três anos.
- Definir a porcentagem de usuários do Outlook que acessam o Exchange Server em cached mode.
- Definir a forma de implementar o HA no servidor Mailbox (LCR, SCC, CCR).
- Definir se vai haver a implementação do DR com o SCR.
- Definir a opção do Backup (VSS clone, Restore LUN, Tape Backup).
- Definir se vai utilizar a opção de antivirus/antispam.
- Definir se vai implementar as outras funções de servidor (Edge, Unified Messaging).
- Definir se vai utilizar os recursos de HA para as outras funções de servidor (Hub Transport Server, Client Access Server Edge Transport Server, Unified Messaging Server).

É necessário considerar o overhead introduzido pelo HYPERVISOR. Por segurança, deve-se considerar 15% de overhead, ou seja, o servidor físico definido pela ferramenta de sizing deve ser 15% maior se o ambiente for virtualizado.

9.8.14. Conceito na Prática: ERP SAP

Introdução

O SAP evoluiu ao longo dos anos e a infraestrutura para suportá-lo ficou ainda mais complexa. Segundo a SAP, 70% das novas soluções vendidas atualmente são baseadas em plataforma aberta. Em grandes instalações é comum toda a infraestrutura de servidores ser baseada na arquitetura x86 e a camada de banco de dados ainda se manter na plataforma RISC ou utilizar um cluster Oracle RAC em plataforma x86. Os sistemas operacionais utilizados são o Windows e o Red Hat Linux. Os bancos de dados utilizados são o MS SQL Server em instalações médias e o Oracle, normalmente utilizado para grandes instalações e muitas vezes baseado no RAC. Os servidores de banco de dados são utilizados em alta disponibilidade. Importante ressaltar que os fabricantes de banco de dados como a Oracle e a IBM mantêm acordos com a SAP para um licenciamento exclusivo.

Fazer o correto *sizing* da arquitetura do SAP, o principal software de gestão empresarial do tipo ERP, é um fator primordial para uma implementação bem sucedida deste tipo de projeto. O *sizing* de um aplicativo do tipo ERP-SAP deve considerar os aspectos de disponibilidade e desempenho (SLA) de um lado e, do outro, o risco e o custo envolvido. Quanto melhor os SLAs, menor o risco e maior o custo do projeto.

A *release* Enterprise 4.7 colocou definitivamente o SAP no mundo web. O SAP NetWeaver incorporou o Unicode Format, que permite que o sistema trate diferentemente aspectos como moeda, taxa, práticas legais em cada país e deu ao SAP aplicabilidade internacional. O servidor de aplicação SAP WEB pode trabalhar com a plataforma proprietária ABAP ou com o J2EE.

A versão atual do SAP ERP é o ECC 6.0, que pode contar com vários módulos FI, CO, MM, PM, FM, HR. Os módulos e a quantidade de usuários definem o sizing do ECC 6.0.

O SAP NetWeaver é uma plataforma de aplicação e integração baseada na web para todas as soluções SAP. A **Tabela 9-11** ilustra a evolução do software SAP com o incremento das suas funcionalidades.

Tabela 9-11 Evolução do SAP

SAP R/3 3.1-4.6	App	SAP basis
SAP R/3 Enterprises	Extensões App CORE	WEB App Server
MySAP ERP	Funções Adicionais Extensões App CORE	SAP NetWeaver (WEB App Server, EP, BI, XI)

SAP NetWeaver

As soluções atuais do SAP são baseadas em SAP NetWeaver, uma plataforma de aplicativos e integração que se propõe a aproveitar a infraestrutura de TI já existente e fornecer um conteúdo de negócios para a camada de integração baseado em diversas indústrias.

O SAP NetWeaver suporta as plataformas ABAP4, JAVA/J2EE e .NET e contém os seguintes elementos:

- Central Instance (processo de message e unqueue).
- DATABASE Instance.
- Dialog Instances (processo dialog, batch,spool, update), se requerido.
- Front-end SAP GUI. As GUIs utilizadas pelo SAP são Windows, Java e HTML.

O SAP NetWeaver pode utilizar uma arquitetura de duas ou três camadas, conforme ilustra a **Figura 9-19**.

- **Duas camadas:** chamada de Central System, onde a instância central e a instância de banco de dados rodam no mesmo servidor.
- **Três camadas:** chamada de Standalone Database System, onde a instância central e a instância de banco de dados rodam em servidores diferentes.

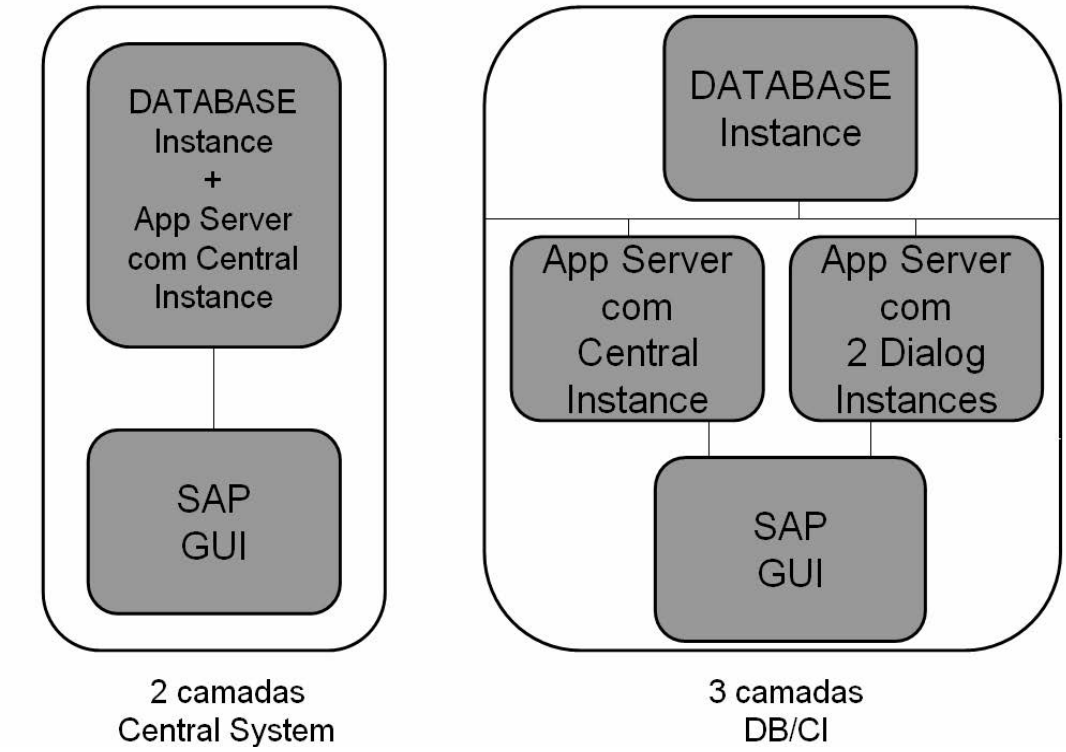

DATABASE Instance + App Server com Central Instance
SAP GUI
2 camadas Central System
DATABASE Instance
App Server com Central Instance
App Server com 2 Dialog Instances
SAP GUI
3 camadas DB/CI

Figura 9-19 SAP em duas e três camadas

Uma arquitetura típica para o SAP é mostrada na **Figura 9-20.** Os servidores de Qualidade e Produção estão em cluster e fazem parte de uma rede SAN baseada em FC. O servidor de desenvolvimento é *stand-alone*. Os servidores que rodam as instâncias Dialog conectam os clientes Windows. Os servidores de produção web *dispatchers* fazem a conexão com os clientes web.

Fazem parte da solução SAP também o Solution Manager (SM), que é uma ferramenta de gerenciamento centralizado para o SAP, e o Solution Landscape Directory (SLD), que é um repositório centralizado de informações sobre os cenários SAP. Estas ferramentas rodam em servidores independentes.

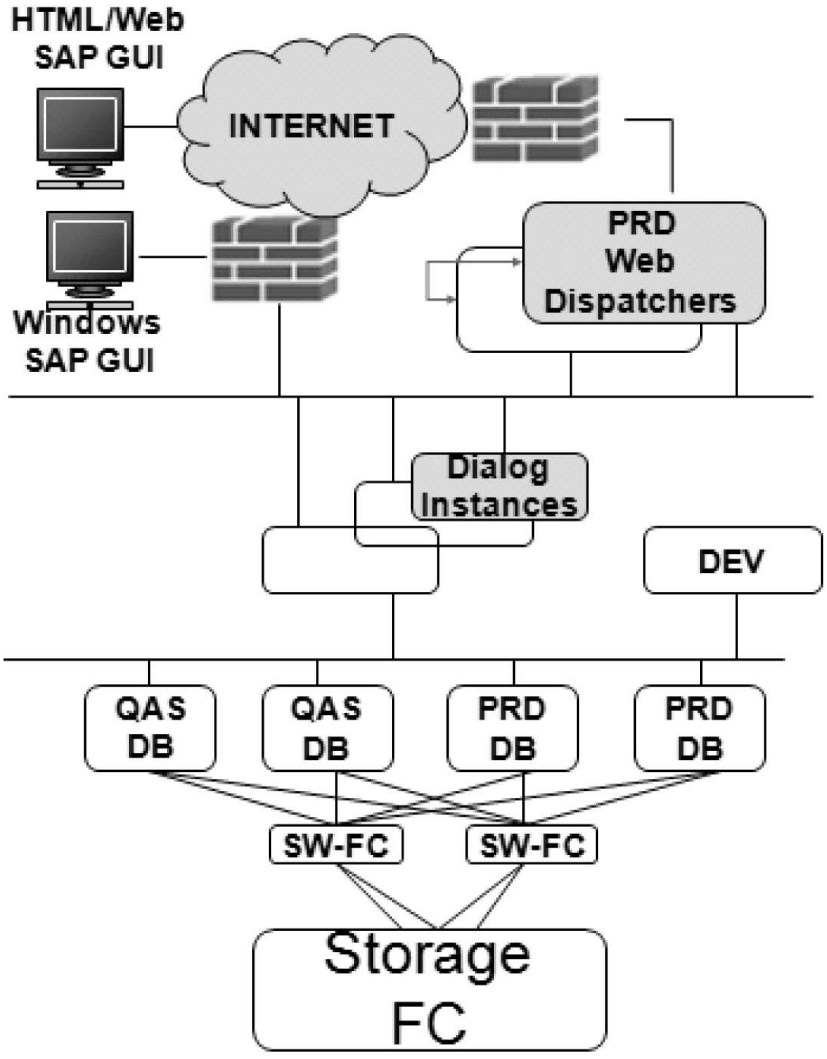


Figura 9-20 Arquitetura para SAP

Cenários SAP

O SAP requer, além da instância de produção (PRD), instâncias adicionais para controle de qualidade (QAS), desenvolvimento (DEV) e, em certas instalações, o treinamento (TRAIN). Estas instâncias são conhecidas como landscapes (cenários) SAP. O que torna ainda mais complexo o sizing do SAP é que, para cada aplicativo (R/3, BI, CRM, Web AS), o SAP requer um cenário próprio. Também, em muitos casos, os cenários do SAP consistem de cópias adicionais dos dados de produção e de testes para funções corporativas. A **Figura 9-21** ilustra os cenários do SAP.

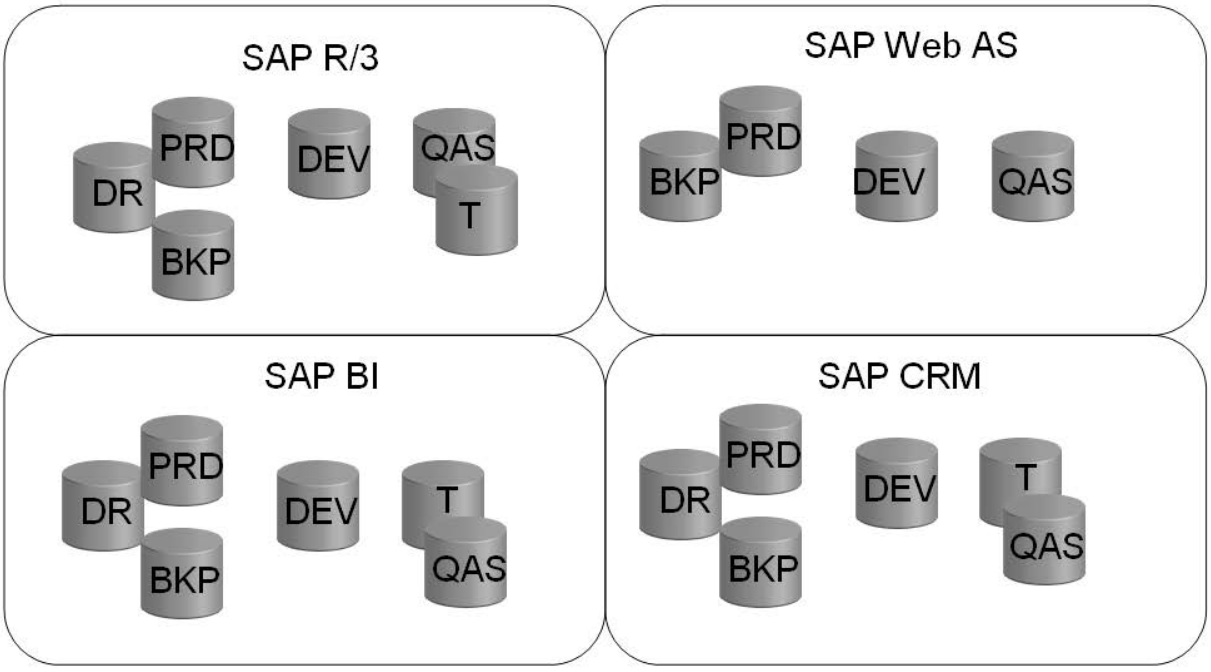


Figura 9-21 Cenários do SAP

Interdependência no SAP

Um outro aspecto relevante do sizing é que, em um ambiente de recuperação de desastres, para manter a consistência dos cenários de negócio, é essencial que todos os cenários sejam reinicializados ao mesmo tempo, o que exige uma tecnologia de replicação sofisticada. Na verdade, uma implementação do tipo SAP *Business Suite* utiliza um banco de dados agrupado. Numa implementação SAP os vários bancos e os vários módulos tornam-se interdependentes.

Os grandes fornecedores mantêm especialistas em sizing de SAP que utilizam software de apoio ao sizing e dados históricos para fazer o correto dimensionamento. Outro aspecto chave é que o SAP é exigente quanto a CPU, memória e I/O, e as bases de dados crescem rapidamente. Portanto, qualquer projeto deve considerar o crescimento esperado, senão o gargalo virá rápido. A camada de banco de dados deve ser cuidadosamente projetada e é necessário verificar os acordos de licenciamento existentes entre

a empresa SAP e os fornecedores de banco de dados como Oracle, Microsoft e IBM. A **Figura 9-22** ilustra a interdependência entre os módulos do SAP.

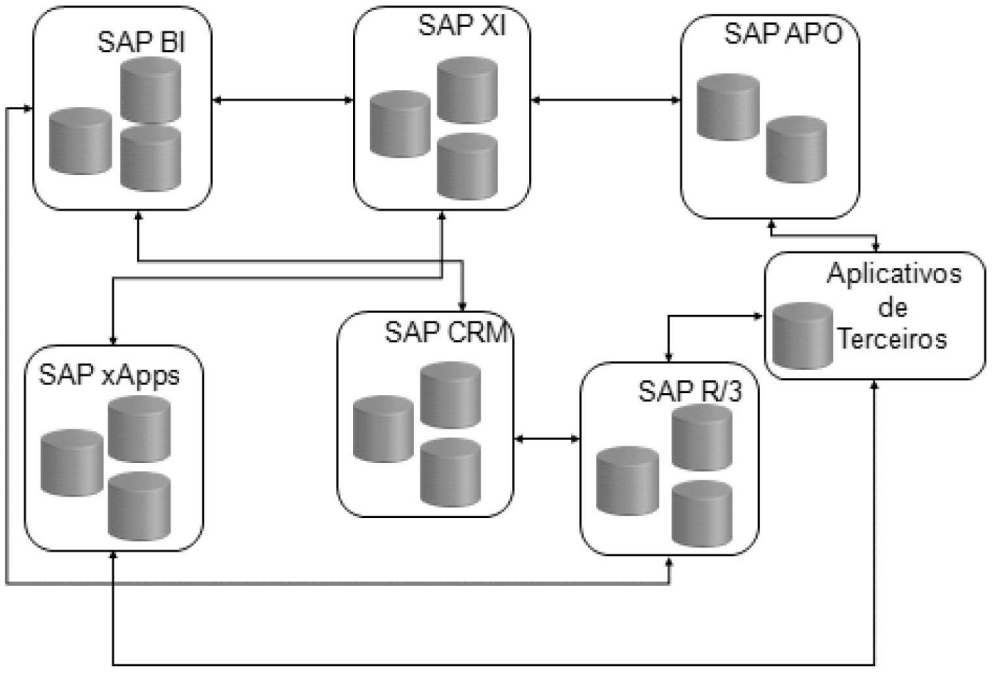


Figura 9-22 Interdependência no SAP

Sizing e Benchmarking SAPS

O SAP Application Performance Standard (SAPS) é um benchmarking independente do hardware que descreve o desempenho de uma configuração de sistema no ambiente SAP. O SAP é derivado do módulo *Sales and Distribuiton* (SD) do SAP, onde 100 SAPS são definidos como 2.000 linhas de itens processados por hora.

O link www.sap.com/solutions/benchmark/sd2tier.epx permite verificar o desempenho das plataformas de hardware para este benchmark. Todo projeto de SAP gera um determinado número de SAPS que define o hardware básico a ser utilizado. Um servidor de 4 processadores-24 núcleos como o R900, da Dell, já consegue uma performance acima de 20.000 SAPS.

O sizing dos diversos módulos do SAP normalmente envolve uma otimização realizada por software. Os SAPS requeridos servem de entrada para softwares que otimizam o hardware considerando aspectos de alta disponibilidade e segurança. Também deve-se sempre estimar um crescimento da base de dados para os próximos três anos (30% pelo menos) e do processamento. A virtualização pode ser utilizada principalmente nos ambientes de teste e qualidade. Inicialmente, em qualquer projeto de SAP, deve-se considerar o sistema operacional a ser utilizado e o sistema gerenciador de banco de dados. Também deve-se considerar o perfil do usuário para cada um dos módulos. Normalmente, empresas que fornecem o hardware para o SAP disponibilizam questionários que permitem identificar as principais necessidades para a realização do sizing.

Outro aspecto vital é a verificação da homologação dos diversos componentes de hardware para a versão do SAP a ser utilizada. Lembrar também que as bases de dados crescem rapidamente e que sistemas como o R/3 e o BI utilizam bases de dados distintas, impactando o tamanho do storage.

A VIRTUALIZAÇÃO do ambiente SAP deve ser pensada com cuidado. Virtualizar o ambiente de teste, desenvolvimento, qualidade e treinamento é quase natural hoje. Virtualizar o ambiente de produção, só se for uma alternativa para melhorar a flexibilidade do ambiente pagando um preço pela perda de performance. Em ambiente Microsoft este penalty é maior do que no ambiente Linux.

A VMware cita as seguintes razões para virtualizar o ambiente SAP:

- Simplificar a transição para ambientes de 64 bitis.
- Facilitar o processo de upgrade para novas versões do SAP.
- Aumentar a disponibilidade para todos os ambientes a um menor custo.
- Criar sistemas de recuperação de desastres com um custo efetivo.
- Desenvolver rapidamente novos ambientes de SAP para desenvolvimento e teste.
- Alinhar o SAP com as prioridades do negócio.
- Diminuir os custos para espaço, energia, refrigeração, hardware e pessoas do DATACENTER SAP.
- Melhorar a utilização do SAP DATACENTER.
- Aumentar o uptime durante manutenção planejada.

O paper *SAP Solutions on VMware Infrastructure: Customer Use Cases da VMware* mostra alguns casos de SAP baseados no ambiente VMware.

9.8.15. Conceito na Prática: Dell Virtualization Advisor Tool

Em adição à ideia de utilizar uma arquitetura de referência, a Dell desenvolveu um *advisor tool* que permite atender necessidades específicas. O Dell Virtualization Advisor Tool, cuja tela principal pode ser vista na **Figura 9-23,** recomenda uma configuração de hardware baseada em dois *inputs*:

- Informação baseada no ambiente existente.
- Informação baseada nas características das WORKLOADS para a nova infraestrutura virtual.

Em ambos os casos, a ferramenta providencia a flexibilidade para escolher as funcionalidades desejadas e a entrada de dados é utilizada para determinar os requisitos de storage, servidores e rede. A saída é uma lista de hardware suportada pela Dell que atende aos requisitos iniciais.

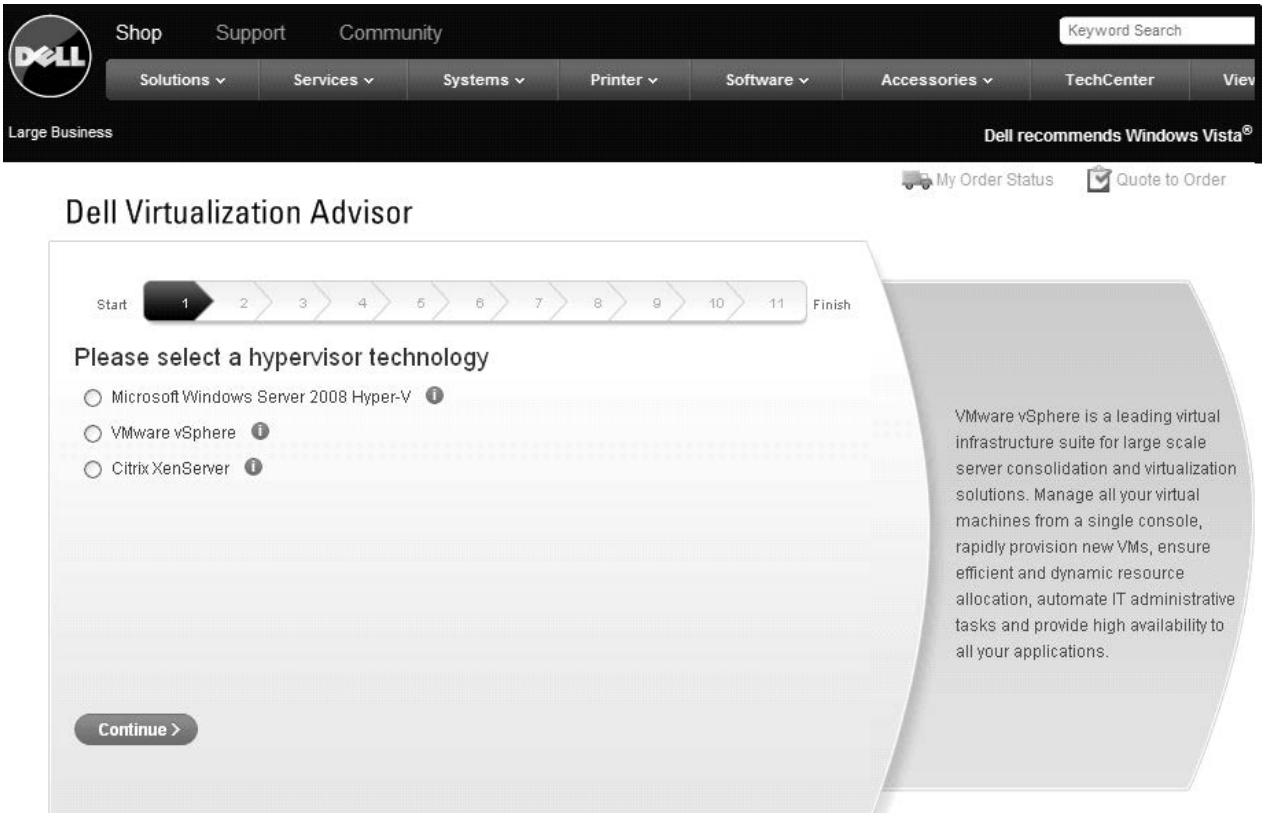


Figura 9-23 Dell Advisor Tool

O **Dell Virtualization Advisor** é uma boa ferramenta para estimar a infraestrutura necessária. Os advisors podem ser utilizados para projetos de infraestrutura existente ou para nova infraestrutura. O link para uso da ferramenta é:

http://advisors.dell.com/AdvisorWeb/Advisor.aspx?advisor=c82c3ec8-c94f-4602-9a41-c20382db1cd0&c=us&l=en&cs=555.

9.9. Implementação

9.9.1. Introdução

A **Figura 9-24** ilustra as etapas a serem vencidas no Projeto 2.

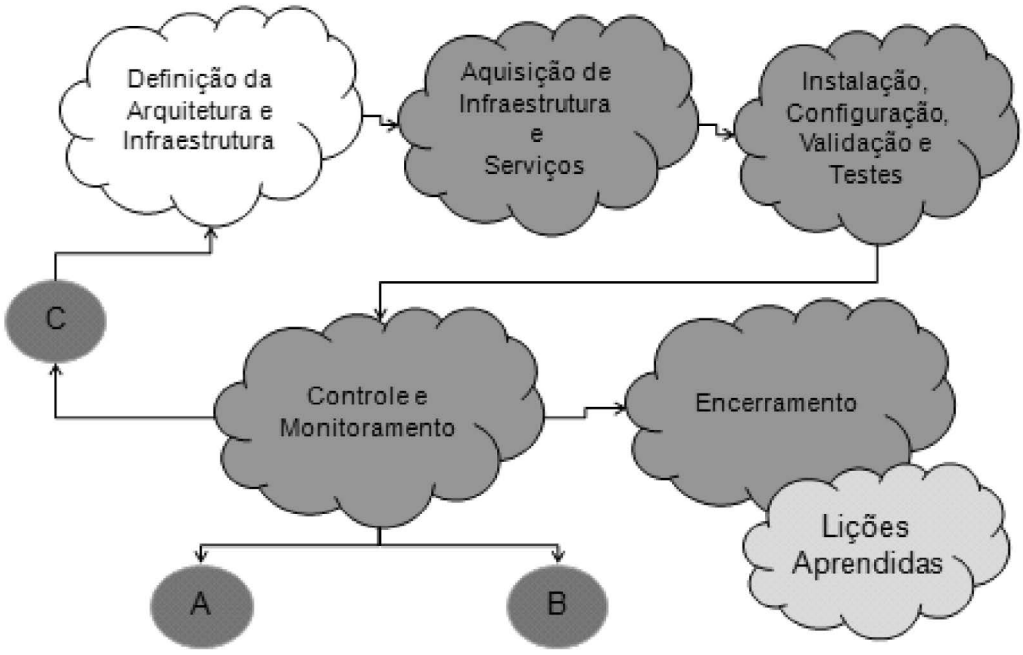


Figura 9-24 Projeto 2 de VIRTUALIZAÇÃO – Etapa 2

9.10. Definição de Arquitetura e Infraestrutura

9.10.1. Introdução

As arquiteturas de referência servem para ajudar a definir a arquitetura para uma infraestrutura que esteja de acordo com os princípios do projeto.

Dicas importantes para servidores:

- Deverão ser utilizados servidores homologados para o software de VIRTUALIZAÇÃO.
- Definição das partições para o servidor virtualizado.
- Definição da forma de licenciamento do software de VIRTUALIZAÇÃO.
- Configurações suportadas para o software de gerenciamento.
- Migração: para converter um servidor físico em servidor virtual (P2V), pode-se utilizar ferramentas dos próprios fabricantes.
- Desenho da Arquitetura deve incluir os servidores, armazenamento e rede.

Dicas importantes para o storage:

- Em geral, o próprio software de VIRTUALIZAÇÃO realiza o failover e o balanceamento de carga.

- Definição da necessidade de definição do uso de zoneamento no storage para efeito de segurança.
- Definição da rede de storage a ser utilizada e funcionalidades necessárias.
- Definição do nível de integração do backup.

Dicas importantes para a contingência:

- A utilização de recursos de alta disponibilidade depende da compatibilidade entre os processadores utilizados nos servidores e do storage utilizado.
- A utilização de recuperação de desastres depende da integração do software de replicação utilizado no storage e a funcionalidade do software de VIRTUALIZAÇÃO.

Princípios básicos de projeto:

- Configuração otimizada do hardware para WORKLOADS virtualizadas.
- Redundância sem ponto de falha.
- Escalabilidade.
- Projeto de rede com isolamento, redundância e alta performance.
- Integração com DATACENTER existente.

9.10.2. Arquiteturas Típicas

Existem muitas escolhas para a arquitetura de um projeto de virtualização. Servidores, storage e software de VIRTUALIZAÇÃO podem variar dependendo dos requisitos de funcionalidade e desempenho requerido.

É normal pensar a arquitetura de referência em três grandes possibilidades: pequena, média e grande. Estas categorias são definidas com base nas capacidades e funcionalidades de cada produto e na arquitetura como um todo. Lógico que qualquer projeto mais detalhado precisa considerar outros aspectos e a necessidade real das muitas funcionalidades.

A categoria mínima, ilustrada na **Figura 9-25,** foca em funcionalidade básica e hardware mínimo. Ela não possibilita obter funcionalidades avançadas que dependem do VMotion da VMware ou do Live Migration da Microsoft.

Uma arquitetura pequena é baseada no ESXi e basicamente possui a opção de rodar as aplicações em um servidor virtualizado de dois processadores, cada processador com quatro cores. O gerenciamento é simplificado e as funcionalidades baseadas no VMotion não são possíveis de serem efetivadas. A configuração pequena, mostrada na figura, utiliza um único servidor e um storage em DAS.

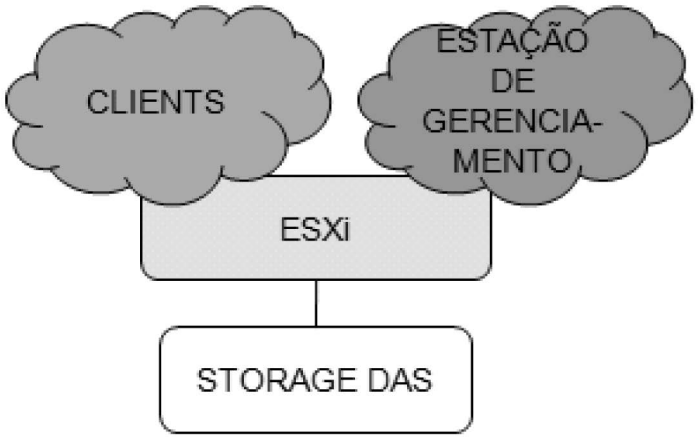


Figura 9-25 Configuração Pequena

A categoria média, ilustrada na **Figura 9-26,** busca a obtenção de requisitos de produção para pequenos e médios negócios com foco em soluções econômicas de SAN.

Uma arquitetura média utiliza o ESXi e apresenta funcionalidades adicionais baseadas no VMotion. Nesta configuração, as funcionalidades HA, FT, DRS e Storage VMotion estão ativadas. A **Figura 9-26** ilustra uma configuração média.

O storage na configuração média está normalmente em uma rede SAN do tipo iSCSI. O Virtual Center roda em um servidor dedicado utilizado também para backup ou pode rodar como máquina virtual. Os servidores ESX são baseados em quatro processadores, cada processador com quatro cores.

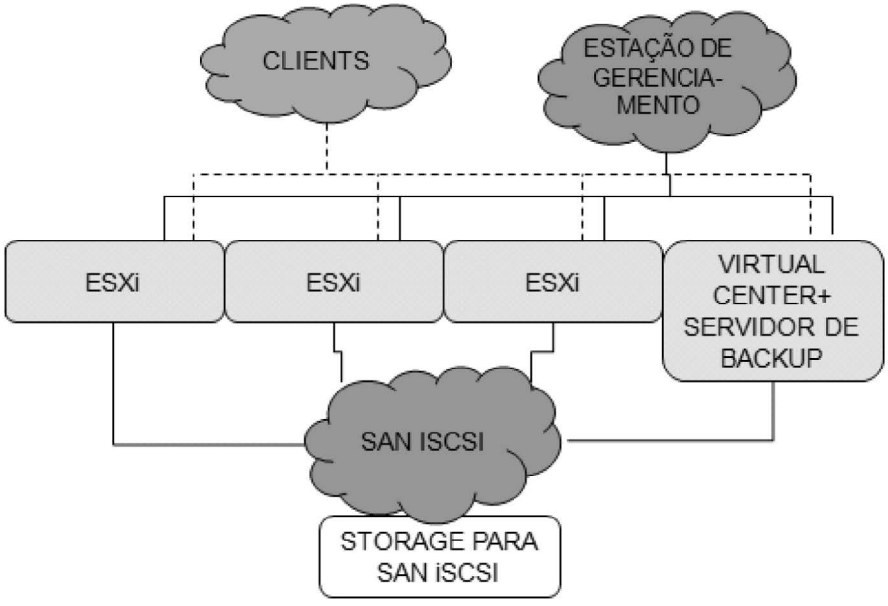


Figura 9-26 Configuração Média

A categoria grande, ilustrada na **Figura 9-27,** foca em soluções de VIRTUALIZAÇÃO completa usando redundância, gerenciamento um para muitos, alta disponibilidade e servidores e storage de classe enterprise.

Uma arquitetura grande utiliza todos os recursos da VIRTUALIZAÇÃO possíveis com o VMware. Além disto, os requisitos de gerenciabilidade e operacionalidade são melhorados.

Os servidores são normalmente do tipo blades. A rede de storage é do tipo SAN FC. Se considerar a recuperação a desastres numa configuração de VIRTUALIZAÇÃO em dois sites, deve-se considerar a integração do software de replicação do storage com a ferramenta SRM.

A VMware sugere que o VIRTUAL CENTER pode ser colocado também em máquina virtual.

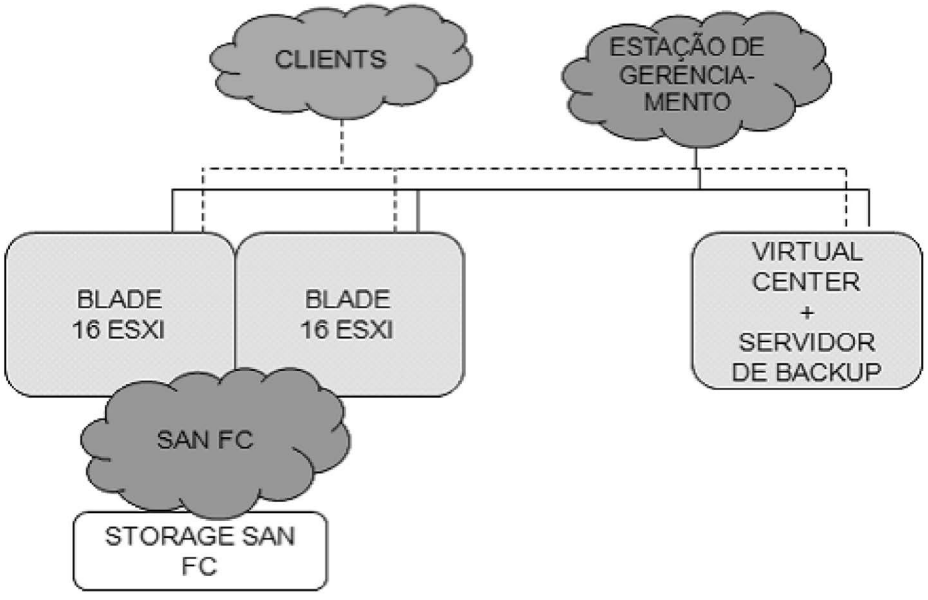


Figura 9-27 Configuração Grande

9.10.3. Conceito na Prática: Dell VMware vSphere

Configurações de referência para Arquitetura de VIRTUALIZAÇÃO utilizada em pequenas e médias empresas possibilitam que organizações que não dispõem de arquitetos de DATACENTER internamente possam rapidamente definir uma configuração que atenda às demandas do negócio. A ideia central é ganhar tempo no projeto ao utilizar uma configuração testada e pronta para uso.

A Dell propõe uma configuração baseada no servidor RACK R510, Storage Equallogic série PS4000 iSCSI e VMware vSphere. O Dell PowerEdge 510 é um servidor para RACK de 2U que suporta o Intel XEON 5600 e memória DDR3. Pode vir com chassi para 4 ou 8 discos. Com 8 discos chega a 8TB de armazenamento interno. O servidor possui naturalmente drivers hot plug, fontes de energia e ventiladores redundantes, além

de duas portas Gigabit Ethernet. O Dell Equallogic PS4000 possui duas portas Gigabit Ethernet por controladora e uma porta 10/100MB de gerenciamento. A Dell também fornece os switches PowerConnect 5424 de 24 portas adequados para conectar o fabric do storage em iSCSI.

A **Figura 9-28** ilustra a arquitetura típica.

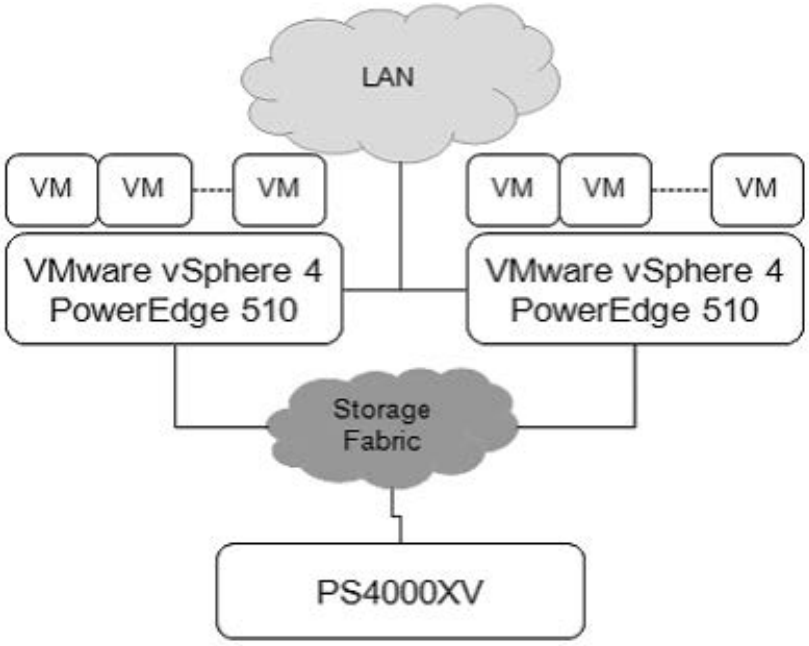


Figura 9-28 Arquitetura de Referência com Equallogic

Os princípios de projeto utilizados para definir uma arquitetura de referência para o SMB são:

- Configuração otimizada para SMB.
- Configuração otimizada dos servidores visando a VIRTUALIZAÇÃO.
- Redundância.
- Arquitetura de rede isolada e redundante.

Baseada nestes princípios, a Dell define três configurações de referência:

- **Stand-alone**: utiliza um servidor R510 com armazenamento interno e o VMware vSphere Essentials Plus. O vCenter pode ser instalado em um servidor PE410 ou em uma máquina virtual.
- **Alta Disponibilidade**: a configuração recomendada consiste de dois servidores R510 rodando o vSphere Essentials Plus conectados ao Equallogic PS4000 em ambiente iSCSI. A infraestrutura de rede consiste de dois switches PowerConnect 5424. O servidor PowerEdge R410 é usado pelo VMware vCenter Server. **A Figura 9-29** ilustra o desenho da solução.

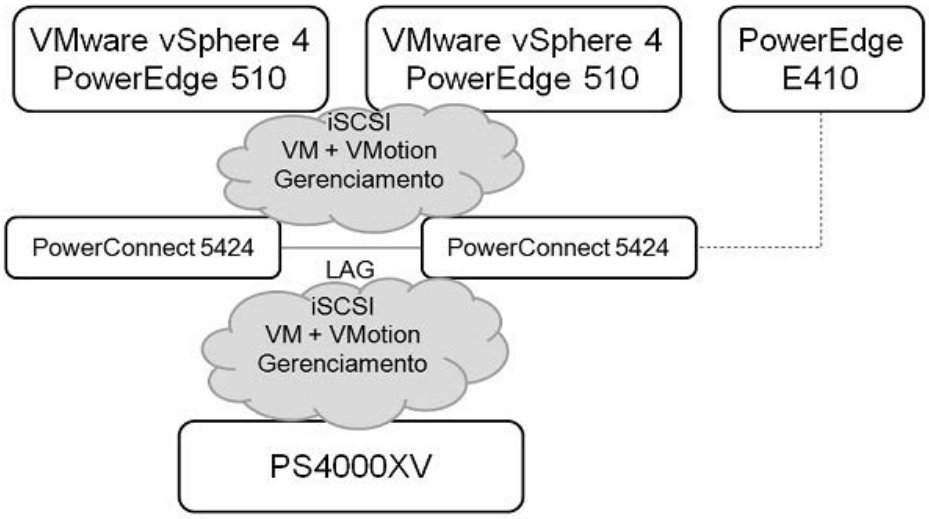


Figura 9-29 Arquitetura de Referência para Configuração HA

- **Alta Disponibilidade + vMotion:** esta configuração requer o VMware vSphere Advanced Edition. A solução consiste de dois servidores PowerEdge 510 rodando o VMware ESX, conectados ao PS4000 em um ambiente iSCSI. Dois switches PowerConnect 5424 são utilizados para suportar a rede de máquinas virtuais, gerenciamento e tráfego VMkernel. Dois switches 5424 são utilizados de forma dedicada ao tráfego iSCSI. O Dell PowerEdge R410 é utilizado para o vCenter Server. A **Figura 9-30** ilustra o desenho da solução.

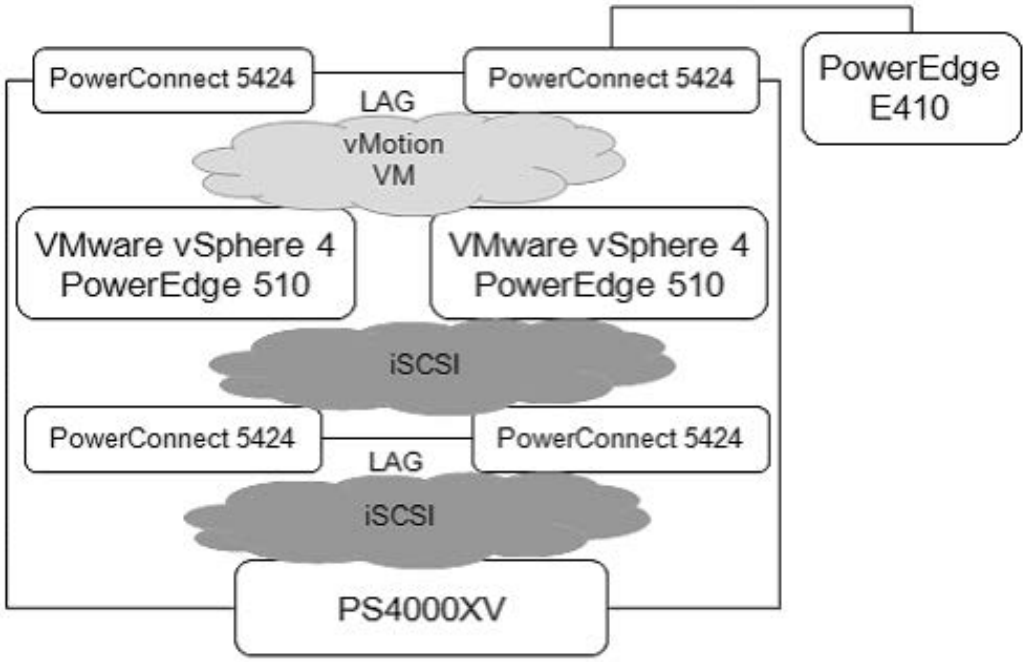


Figura 9-30 Arquitetura de Referência para Configuração HA + vMotion

A infraestrutura do VMware vSphere compreende quatro tipos de tráfego:

- Tráfego das máquinas virtuais.
- Tráfego do gerenciamento.
- Tráfego do vMotion.
- Tráfego iSCSI.

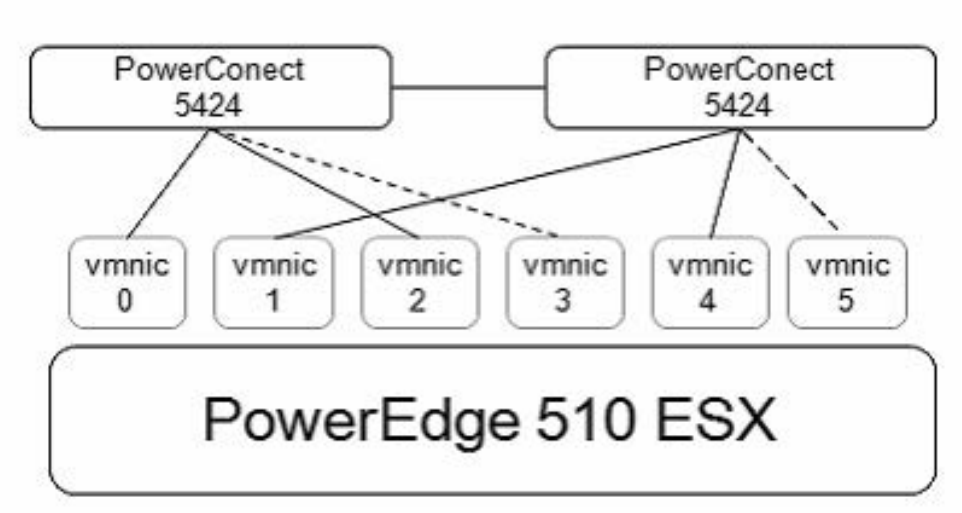
As configurações HA e HA + vMotion requerem que cada host tenha 6 portas de rede (duas NICS integradas e dois adaptadores dual port, totalizando 6 portas). Para cada R510 usado na configuração, os adaptadores físicos são enumerados da seguinte forma:

- vmnic 0, vmnic 1: LOMs integrados na placa do servidor.
- vmnic 2, vmnic 3: primeira e segunda porta na primeira add-on dual port Broadcom.
- vmnic 4, vmnic 5: primeira e segunda porta na segunda add-on dual port Broaddcom.

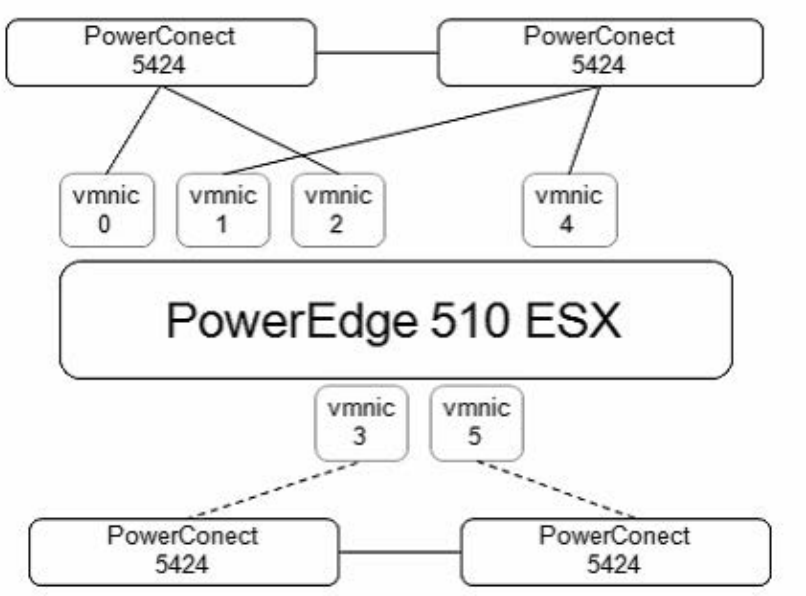
Os switches 5424 são interligados usando um LAG (link agregation) de 4 portas.

- Na configuração HA todos os tráfegos de rede usam os dois switches que são isolados usando VLANs.
- Na configuração HA + VMotion são utilizados 4 switches, dos quais dois são isolados para a rede iSCSI e as outros dois para o tráfego das VMs, a rede de gerenciamento e a rede vMotion.

A **Figura 9-31** (a) e (b) ilustra as duas configurações de rede.



(a)



(b)

Figura 9-31 Configuração de rede para configurações HA e HA + vMotion

Os switches virtuais são criados pelo VMware. O vSwitch 0 é criado utilizando 4 adaptadores de rede (vmnic4,vmnic2, vmnic1, vmnic0) como uplinks. O vSwitch 1 é criado e vmnic 3 é adicionado a este switch. O vSwitch 2 é criado e o vmnic 5 é adicionado a este switch. No caso da configuração para HA, os tráfegos precisam ser separados pelas VLANs. No caso HA + vMotion não é necessário, pois utilizam-se quatro switches físicos.

9.10.4. Conceito na Prática: Cisco EMC VMware vSphere Vblock

VMware, Cisco e EMC estão juntas em uma parceria chamada de VCE (*Virtual Computing Environment*) para oferecer uma solução para nuvens privadas baseadas no conceito de Vblock. Os Vblocks são pacotes de infraestrutura otimizados para a construção da nuvem privada.

Estas estruturas integradas facilitam a construção dos grandes blocos da nuvem privada e ofertam de maneira integrada: VIRTUALIZAÇÃO, processamento, armazenamento, rede, segurança e gerenciamento integrado.

Os pacotes Vblock são pré-testados e validados, com desempenho, capacidade e disponibilidade definidos. A promessa é disponibilizar um caminho acelerado para a formação de nuvens privadas.

Os pacotes Vblock surgiram de uma ideia de simplificar a aquisição, a distribuição e a operação da infraestrutura de TI. O valor da solução encontra-se em uma combinação da eficiência, controle e escolha. Outro princípio que guia a ideia dos pacotes de infrastrutura como o Vblock é a habilidade de expandir a capacidade dos pacotes porque a arquitetura é muito flexível.

Os pacotes Vblock são otimizados para fornecimento de SLAs padrões, reduzem o risco e atendem a necessidades de compliance. Para os clientes, simplificam a expansão e a escalabilidade, permitem adicionar storage e capacidade quando requerido, podem se conectar à infraestrutura existente de LAN e permitem administração multitenant e segurança baseada em regras.

Os pacotes Vblock definem as necessidades de espaço e peso, energia e refrigeração e predeterminam a escalabilidade e capacidade do DATACENTER.

Pacotes de Solução

São três os pacotes do tipo Vblock:

- **Vblock 2.** Projetado para instalações high-end. Envolve Cisco Unified Computing System, EMC Symmmetrix VMAX storage e VMware vSphere 4.
- **Vblock 1.** Projetado para grandes instalações. Envolve Cisco Unified Computing System, EMC Clariion CX4 ou EMC Celerra Unified Storage e VMware vSphere.
- **Vblock 0.** Projetado para sites moderados e grande número de máquinas virtuais. Envolve Cisco Unified Computing System, EMC Celerra Unified Storage e VMware vSphere 4.

A **Tabela 9-12** resume as características principais dos três blocos.

Tabela 9-12 Configurações do Vblock

	VBLOCK 0	VBLOCK 1	VBLOCK 2
HYPERVISOR	vSphere 4 ESXi	vSphere 4 ESX	vSphere 4 ESX
Método de BOOT	Local	SAN	SAN
vCENTER	SIM	SIM	SIM
Nexus 1000v	SIM	SIM	SIM
UCB B-series	UCB B-200/250	UCB B-200/250	UCB B-200/250
UCB Mínimo	4*B200, Intel X5550, 48GB RAM	12*B200, Intel X5570, 48GB RAM	32*B200, Intel X5570, 96GB RAM
UCB Máximo	16*B200, Intel X5550, 48GB RAM	24*B200, Intel X5570, 48GB RAM	64*B200, Intel X5570, 96GB RAM
CNA (FCoE)	Emulex, Qlogic, VIC	Emulex, Qlogic, VIC	Emulex, Qlogic, VIC
VMs	128-512	512-3000	3000-20000
UCS Manager	SIM	SIM	SIM

	VBLOCK 0	VBLOCK 1	VBLOCK 2
Storage	CELERRA NS-120	CLARIION CX4-480	SYMMMETRIX V-MAX
Storage Mínimo	5*600GB 15K 5*1TB SATA	6*400GB EFD 78*450GB 15K 21*1TB SATA	9*400GB EFD 124*450GB 15K 76*1TB SATA
Storage Máximo	25*600GB 15K 15*1TB SATA	16*400GB EFD 78*450GB 15K 146*450GB 15K	21*1TB 27*1TB SATA
NAS GATEWAY	N/A	NS-G2 (recomendado)	NS-G8 (recomendado)
Switch SAN	N/A	MDS 9506 (preferido)	MDS 9506 (preferido)
UIM	Preferível	Preferível	Preferível

9.10.5. Vblock 0

O Vblock 0 é uma plataforma ideal para companhias que buscam uma configuração de nível de entrada para o seu DATACENTER. Inclui basicamente Cisco UCS, EMC Celerra e VMware vSphere 4.

Componentes:

- Cisco UCS-B-Series half-slot Blade Servers 4-16 (3-15 servidores blade para a WORKLOAD e 1 para gerenciamento).
- Cisco Nexus 1000v.
- UCS 6100 series fabric.
- Celerra NS-120 single ou dual em um único RACK, indo até 46TB.
- VMware vSphere com ESXi 4.0.
- EMC Ionix UIM (opcional).
- VMware vCenter 4.0.
- EMC PowerPath/VE.
- Cisco UCS Manager.

A **Figura 9-32** ilustra a topologia de alto nível do Vblock 0.

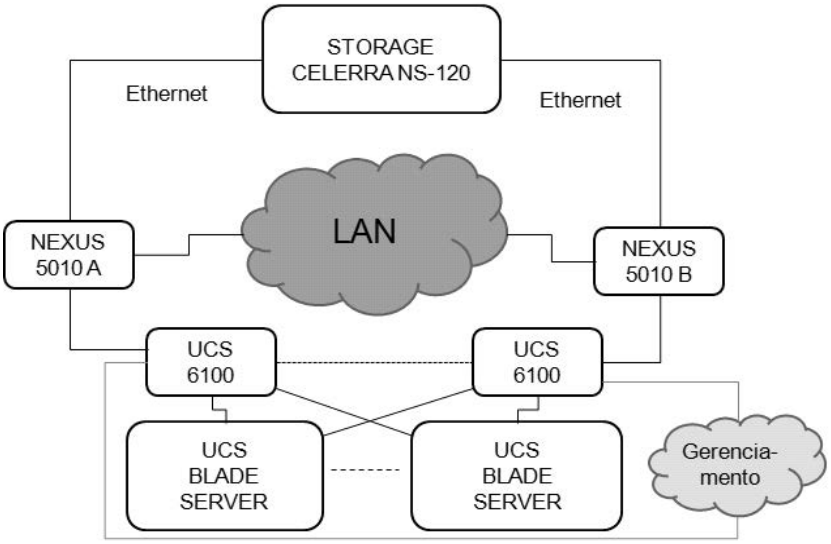


Figura 9-32 Topologia de Alto Nível – Vblock 0

A Figura 9-33 ilustra o RACK Vblock 0**.**

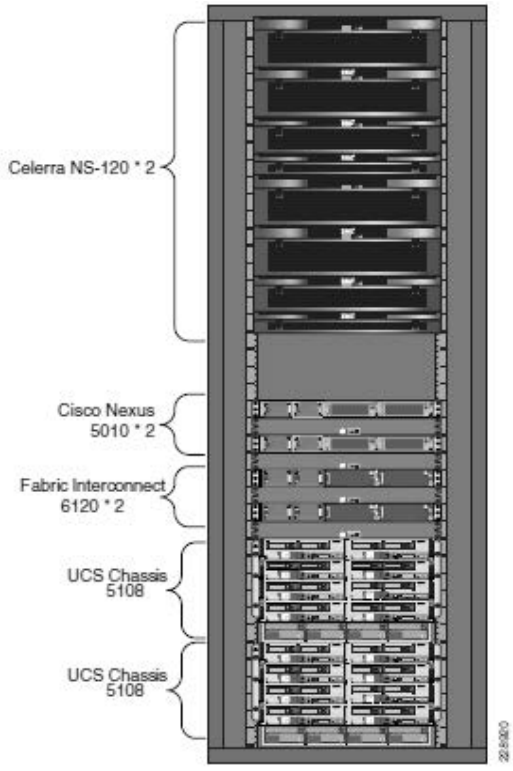


Figura 9-33 Vblock RACK Layout

9.10.6. Conceito na Prática: EMC Ionix

O UIM (*Unified Infrastructure Manager*, gerenciador de infraestrutura unificada) do EMC Ionix permite provisionamento simplificado e integrado, gerenciamento de configuração, alteração e conformidade para VIPs (*Vblock Infrastructure Packages*, pacotes de infraestrutura Vblock).

Como o mais recente acréscimo ao portfólio de gerenciamento de TI do EMC Ionix, o UIM do Ionix é criado especificamente para o gerenciamento de elementos da infraestrutura Vblock, gerenciando os Vblocks como uma só entidade. Isso reduz as despesas operacionais – enquanto fornece recursos de gerenciamento simplificado que capacitam as empresas a fazer mais facilmente a transição de infraestruturas físicas para virtuais e para nuvem privada. Na versão inicial, o UIM do Ionix oferecerá gerenciamento de suporte total do UCS (*Unified Computing System*, sistema de computação unificada) da Cisco. Em seguida, evoluirá para um suporte completo de infraestruturas Vblock.

Principais recursos

- **Painel de controle Vblock e portal de provisionamento de TI**: com o painel de controle UIM do Ionix, é possível visualizar diversas implantações de Vblock, dando a você exibições consolidadas de toda a sua infraestrutura Vblock.
- **Catálogo dos perfis de serviços**: perfis de serviços são a "receita" para a criação de serviços e a base para a entrega da "infraestrutura como serviço". Com o UIM do Ionix, é possível criar vários níveis de perfis de serviços para configurações de rede, computação e armazenamento e combiná-los em uma "oferta" para os clientes. Isso habilita os componentes Vblock a serem fornecidos como unidades padrão de serviço para um cliente, selecionado por meio do painel de controle Vblock. Os recursos do perfil de serviços serão fornecidos para a infraestrutura Cisco na versão inicial.
- **Gerenciamento baseado em políticas**: as políticas de configuração podem ser definidas e aplicadas para garantir a conformidade em todo o sistema, a fim de evitar um fluxo de configurações. A primeira versão do UIM do Ionix oferece esses recursos para o UCS da Cisco e para a infraestrutura de rede relacionada, como o Cisco Nexus e dispositivos MDS.
- **Provisionamento unificado, configuração e alteração**: o UIM do Ionix oferece uma implantação de infraestrutura automatizada e provisionamento de hardware vazio. Isso inclui um provisionamento automatizado de infraestrutura de local para recuperação de desastres. O UIM aproveita o contexto entre domínios para gerenciar dependências entre software, hardware, rede e armazenamento. A detecção profunda e granular oferece uma base para o provisionamento, a configuração e a alteração do UIM do Ionix – permitindo um histórico de revisão ilimitado e rastreamento individualizado, capacidade de rastreamento e de reprodução. A primeira versão do UIM do Ionix concentra-se no provisionamento unificado, na configuração e na alteração para o UCS da Cisco e para a infraestrutura de rede relacionada, como o Cisco Nexus e dispositivos MDS.

- **Integração simplificada**: o gerenciamento de elementos Vblock UIM do Ionix integra-se a sua solução de gerenciamento corporativo, oferecendo eventos de alteração e conformidade para ajudá-lo a rastrear alterações essenciais. O UIM também aproveita as APIs do UCS Manager da Cisco e os sistemas de gerenciamento de armazenamento da EMC, obtendo os benefícios da instrumentação existente e da configuração de componentes.

A **Figura 9-34** ilustra a plataforma EMC Ionix. Eventos de configuração e conformidade e eventos de disponibilidade e desempenho alimentam a plataforma de gerenciamento corporativo.

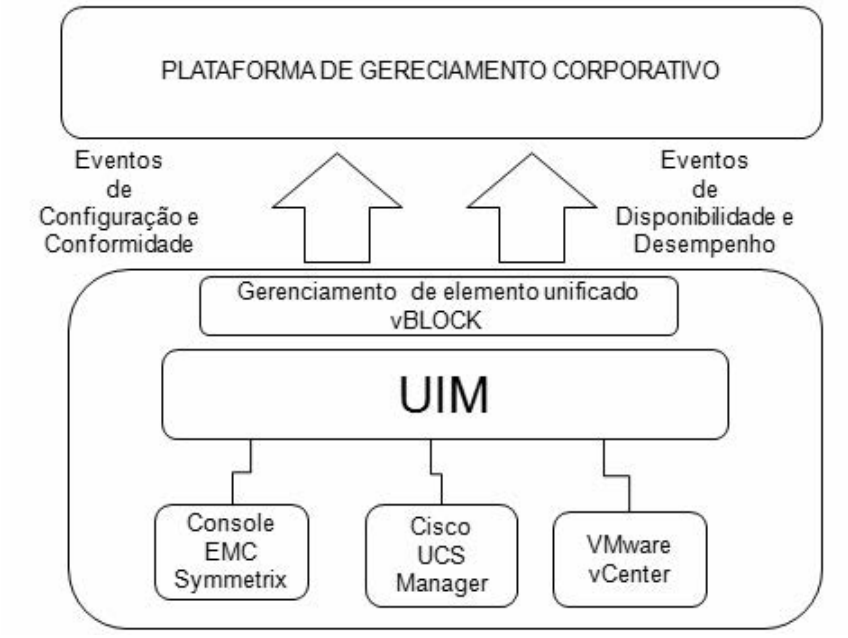


Figura 9-34 EMC Ionix

9.10.7. Aquisição de Infraestrutura e Serviços

Chegou a hora de adquirir o hardware necessário, o software de virtualização e os serviços de terceiros. Normalmente os fornecedores principais de hardware também possuem equipes especializadas na instalação e configuração do software de VIRTUALIZAÇÃO e da ferramenta de gerenciamento.

9.10.8. Instalação, Configuração, Testes e Validação

Deve-se instalar a nova infraestrutura e testar e validar a solução. A validação trata de demonstrar que os requisitos do projeto foram atendidos. Os testes validam os requisitos e definem se o projeto atendeu as expectativas.

9.11. Monitoramento e Controle

9.11.1. Definição dos Níveis de Serviço e Indicadores

Deve-se definir os níveis de serviço e a forma de monitorá-los quando da operação da nova infraestrutura.

9.12. Encerramento

Deve-se encerrar formalmente o projeto encerrando os contratos. Documentar as lições aprendidas é tarefa importante também.

9.12.1. Lições Aprendidas

Deve-se documentar as lições aprendidas ao final do projeto. Em grandes empresas este é um aspecto vital e significa otimizar novos projetos.

9.13. Referências Bibliográficas

Building the Virtualized Enterprise, VMware white paper, 2006.

Capacity Planning: The Blueprint of Server Consolidation, VMware, 2006.

Gerenciamento de infraestrutura unificada do EMC Ionix, DATA SHEET, EMC, 2010.

Reducing Server Total Cost of Ownership with VMware Virtualization Software, VMware, 2006.

The True Cost of Virtual Server Solutions, Taneja Group, 2009.

Vblock 0 Reference Architecture, Cisco EMC VMware, 2010.

Veras, Manoel. DATACENTER: Componente Central da Infraestrutura de TI, Brasport, 2009.

Virtual Computing Environment Coalition Vblock Infrastrcture Packages, Cisco EMC VMware, 2010.

Virtualization Reality: Why Microsoft Virtualization Solutions Deliver Value when Compared to VMware, Microsoft, 2009.

Why Choose VMware?. White Paper, VMware, 2010.

VMware vSphere Reference Architecture for Small Medium Business. Dell Virtualization Business Ready Configuration. Dell, 2009.

10. Case vSphere

10.1. Introdução

Este caso trata de um projeto de VIRTUALIZAÇÃO desenvolvido pela Aliança Consultoria junto à SEFAZ do Rio Grande do Norte. A metodologia utilizada é a que desenvolvo no capítulo 9.

10.2. Business Case, Escolha do Fornecedor e Versão do Software

A justificativa do projeto internamente foi conduzida pelo cliente utilizando os argumentos e ferramenta de TCO-ROI, mostrada no Capítulo 9. A escolha do fornecedor do software de VIRTUALIZAÇÃO foi baseada nos requisitos iniciais do projeto com ênfase nos níveis de serviço que deveriam ser obtidos para as aplicações (R07 da documentação de requisitos).

A **Figura 10-1** ilustra esta etapa.

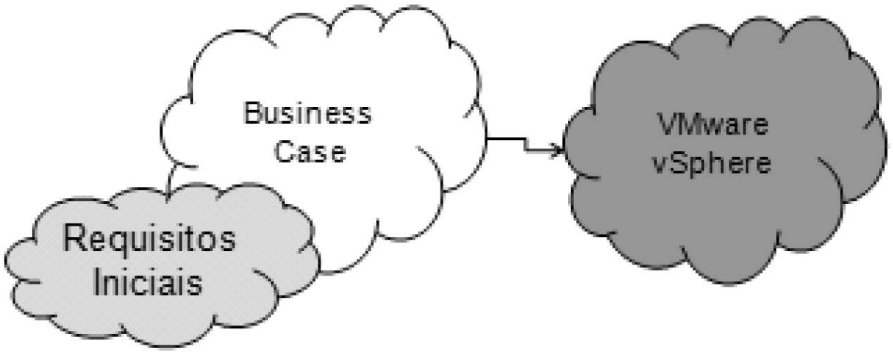


Figura 10-1 Definição do Fornecedor de Software

10.3. Documentação de Requisitos

A **Tabela 10-1** ilustra os requisitos do projeto. O software escolhido foi o VMware vSphere 4.0 em sua versão Enterprise Plus. Esta escolha de versão foi baseada nos requisitos necessários do projeto.

Tabela 10-1 Requisitos

Número	Descrição
R01A/B	O DATACENTER de produção deve ser compatível com a filosofia de TI verde, ou seja, considerar a redução do consumo de energia elétrica, principalmente nos períodos de pouca atividade (como, por exemplo, à noite). O objetivo é reduzir o custo operacional do DATACENTER. O DATACENTER de produção deve atingir os objetivos atuais de CAPACIDADE e prover maior ESCALABILIDADE prevendo demandas futuras.
R02	O ambiente de produção deve ser isolado logicamente do ambiente de desenvolvimento e teste.
R03	Os servidores na DMZ devem estar logicamente isolados em termos de conectividade (VLAN) dos demais servidores internos.
R04	Os servidores de banco de dados Oracle atuais serão migrados para o cluster da Oracle, o Real Application Cluster, que permitirá obter alta disponibilidade e aumento da capacidade de processamento. Desta forma, os servidores não serão virtualizados.
R05	O sistema de cópia de segurança deve ser confiável, rápido e abranger todos os dados do DATACENTER.
R06	As proteções operacionais e de desastres devem permitir recuperação contra falhas lógicas sem perda de dados, ou seja, permitir recuperação rápida dos dados. Exemplos de falhas: falhas humanas (exclusão acidental) e corrupções lógicas de arquivos.
	DISPONIBILIDADE E DESEMPENHO: SLA para as Aplicações
R07	O novo DATACENTER deve atender aos níveis de serviço (SLA) expostos na tabela de catálogo de nível de serviço para as seguintes aplicações/níveis: • **NÍVEL 1** – SISTEMA PRINCIPAL, NFE, SPED E SERVIÇOS OFERECIDOS ON-LINE PARA CONTRIBUINTES. • **NÍVEL 2** – E-MAIL, SERVIÇO DE DIRETÓRIO, DATA WAREHOUSE E SISTEMAS LEGADOS. • **NÍVEL 3** – DEMAIS APLICAÇÕES DE PRODUÇÃO. • **NÍVEL 4** – DESENVOLVIMENTO E HOMOLOGAÇÃO.

10.4. Declaração do Escopo

Consolidação e VIRTUALIZAÇÃO do DATACENTER principal de produção/ desenvolvimento/ homologação que considere aspectos de redução de custos de energia e refrigeração e escalabilidade.

Implantação de um DATACENTER remoto para uso em caso de falha e contenção na infraestrutura principal e implantação de infraestrutura de cópia de segurança e recuperação das informações.

10.4.1. Objetivo Geral

Desenvolver um ambiente de TI que garanta alta disponibilidade e continuidade dos negócios mesmo em caso de falhas leves ou graves (desastre).

10.4.2. Objetivos Específicos

Os objetivos foram colocados em termos de disponibilidade, capacidade e custos.

- Alcançar uma disponibilidade de 99,9% (equivalente a 8:45 de downtime em um ano) nos seus serviços de nível 1.
- Alcançar uma disponibilidade de 99% (equivalente a 3 dias e 15:36 de downtime em um ano) nos seus serviços de nível 2.
- Entregar uma infraestrutura inicial com capacidade total equivalente a 130% do ambiente atual em termos de recursos computacionais e de armazenamento.
- Entregar uma nova infraestrutura com alta escalabilidade. Deve haver a garantia de que o novo ambiente esteja pronto para crescer em torno de 300% em 5 anos.
- Reduzir o custo operacional do DATACENTER principal em 25% nos doze meses subsequentes ao início da operação.

10.4.3. Pressupostos

Os pressupostos do projeto são colocados na **Tabela 10-2**.

Tabela 10-2 Pressupostos do Projeto

Número	Descrição
P01	A sala do novo DATACENTER possui infraestrutura física adequada de conectividade, cabeamento, energização e refrigeração.
P02	Existe conectividade IP entre os DATACENTERS primário e de contingência para replicação de dados e comunicação de LAN. A qualidade deste link atende aos requisitos de SLA expostos neste documento do ponto de vista de latência e largura de banda.

P03	A instituição dispõe de software, equipamentos, appliances e equipe técnica com capacidade para adaptar a solução de segurança da informação ao novo ambiente.
P04	Há conectividade WAN redundante para acesso dos usuários via internet com os DATACENTERS de produção e contingência.

10.4.4. Demarcações e Regras

Tabela 10-3 Demarcações e Regras

Número	Descrição
D01	Os servidores físicos atuais de produção, blades HP com lâminas DL 460c, serão utilizados no contigenciamento dos serviços no segundo DATACENTER.

10.4.5. Identificação de Riscos

- Ausência de software de automatização e teste da política de contingência de DATACENTER (*VMware Site Recovery Manager*) pode inviabilizar o cumprimento do SLA para aplicações de nível 1.
- Algumas aplicações em sistemas operacionais não Microsoft Windows podem não ser migradas, através de software de migração P2V, da plataforma física para a virtual, exigindo que seja feita uma nova instalação do servidor diretamente na infraestrutura virtual.

10.5. Situação Atual

10.5.1. Inventário

Aplicações

- Microsoft SQL Server
- Microsoft Active Diretory
- Microsoft Internet Information Services
- Apache / Tomcat
- Oracle Application Server
- Squid Proxy Server
- Servidor de arquivos

Servidores

Foram retirados do inventário todos os servidores que hospedam banco de dados Oracle. A **Tabela 10-4** ilustra o inventário de servidores.

Tabela 10-4 Servidores

Descrição	Quantidade
Total de servidores (físicos e virtuais)	35
Total de servidores físicos	25
Total de servidores de produção	30
Total de servidores de desenvolvimento/homologação	5
Total de servidores Windows	28
Total de servidores Linux	4
Total de servidores SunOS	3

Armazenamento

O armazenamento encontrado foi do tipo DAS e SAN, conforme descrito:

- Armazenamento local (DAS) em 60% dos servidores.
- Armazenamento compartilhado via SAN Fibre Channel em storage HP EVA 4100 com discos FATA de 500 GB, totalizando 28 discos.
- O baixo desempenho do storage, principalmente pelo tipo de disco adotado, compromete o seu uso atual para as aplicações que requerem grande quantidade de I/O, notadamente banco de dados relacionais.

Rede

- A grande maioria dos servidores (cerca de 85%) utiliza conectividade Gigabit Ethernet através de switch de borda layer 2 tipo RACK e blade.

10.5.2. Análise da Demanda das WORKLOADS

As análises descritas a seguir foram baseadas em dados obtidos por utilitários de coleta de desempenho. Os servidores Oracle foram excluídos deste diagnóstico.

A **Tabela 10-5** ilustra o estado atual da CPU.

Tabela 10-5 Estado Atual de CPU

Estado atual de CPU	
Métrica	**Quantidade**
Número médio de CPUs por servidor físico	3,3
Média de Mhz por CPU	2.350 Mhz
Quantidade média de Mhz por servidor físico	7.755 Mhz
Média de utilização de CPU por servidor físico	40% (3.100 Mhz)
Total de recurso de CPU requerido atualmente no DATA-CENTER	77.500 Mhz

Tipos de processadores encontrados:

- 65% Intel Xeon E5320 e E5335
- 17,5% Intel Pentium 4
- 17,5% Intel Core 2

A **Tabela 10-6** ilustra o estado atual da RAM.

Tabela 10-6 Estado Atual de RAM

Estado atual de RAM	
Métrica	**Quantidade**
Soma total de RAM dos servidores físicos	109.000 MB
Média de utilização de RAM por servidor físico	80%
Total de recurso de RAM requerido atualmente no DATA-CENTER	87.200 MB
Média de RAM utilizada por servidor físico	3.488 MB

Tipos de memória encontrada:

- Todas as memórias são DDR2.
- A utilização da memória nos servidores varia entre 512 MB a 8196 MB.

A **Tabela 10-7** ilustra o estado atual do armazenamento.

Tabela 10-7 Estado Atual de Armazenamento

Estado atual de armazenamento	
Métrica	**Quantidade**
Capacidade total de armazenamento para produção/dev/homolog	7 TB líquidos
Banda de conectividade host-storage	4 Gb/s
Capacidade total de IO/s de armazenamento de produção (4 KB)	Até 2.500 IOPS
Média de utilização de espaço para produção	44%
Total de espaço de armazenamento requerido atualmente	3,5 TB

10.6. Definição da Arquitetura e da Infraestrutura

10.6.1. Especificações Gerais do Projeto

E01

A **Tabela 10-8** ilustra o requisito R01 e a especificação E01.

Tabela 10-8 Requisito e Especificação 01

Requisito:	R01A/B
	O DATACENTER de produção deve ser compatível com a filosofia de TI verde. O DATACENTER de produção deve atingir os objetivos atuais de CAPACIDADE e prover maior ESCALABILIDADE prevendo demandas futuras.
Especificação:	E01
	Máxima consolidação do DATACENTER, reduzindo a quantidade de equipamentos físicos como servidores, cabos, ativos, RACKS, etc. Esta redução promove um menor consumo de energia elétrica para refrigeração e alimentação dos componentes. Utilização de equipamentos que possibilitem o atingimento da capacidade requerida e a escalabilidade do ambiente para os próximos cinco anos.

- A fim de buscar uma alta taxa de consolidação de servidores virtuais por servidor físico, contribuindo assim para a consolidação máxima do ambiente, utilizar as diversas características de gerenciamento de recursos e concorrência (controle de nível de serviço).
  - Gerência de processador e memória – VMware vSphere 4 Resource Pool.
  - Gerência de I/O de armazenamento – EMC Clariion Navisphere Quality of Service.
- Otimização da consolidação física com substituição de servidores RACK e torre por sistema blade para maximizar o consumo elétrico e a redução da área física do CPD.
- Redução extra de energia com o desligamento (stand by) automático de hosts físicos (e a devida migração das máquinas virtuais neste host para os hosts restantes) em momentos de pouca atividade do DATACENTER, através do recurso vSphere 4 Distributed Power Management (DPM).

E02

A **Tabela 10-10** ilustra o requisito R02 e a especificação E02.

Tabela 10-9 Requisito e Especificação 02

Requisito:	R02
	O ambiente de produção deve ser isolado logicamente do ambiente de desenvolvimento e produção.
Especificação:	E02

	- Compartilhar os mesmos servidores físicos (cluster) e sistema de armazenamento, contudo:
	  • Diferenciando o desempenho do ambiente de desenvolvimento, a fim de que este não atrapalhe o SLA da produção, através do VMware vShpere 4 Resource Pool.
	  • Isolando o gerenciamento e o acesso dos ambientes para diferentes operadores e administradores, através de regras de permissão aplicadas ao Resource Pool.

- Exemplo:

E03

A **Tabela 10-10** ilustra o requisito R03 e a especificação E03.

Tabela 10-10 Requisito e Especificação 03

Requisito:	R03
	Os servidores na DMZ devem estar logicamente isolados em termos de conectividade (VLAN) dos demais servidores internos.
Especificação:	E03
	Utilizar Private Virtual LANs (PVLAN) para permitir isolamento layer 2 entre servidores internos e DMZ na mesma rede, através de vSphere 4 Distributed Virtual Switches.

E04

A **Tabela 10-11** ilustra o requisito R04 e a especificação E04.

Tabela 10-11 Requisito e Especificação 04

Requisito:	R04
	Os servidores de banco de dados Oracle atuais serão migrados para servidores físicos com o Real Application Cluster (RAC).
Especificação:	E04
	Os bancos de dados Oracle serão distribuídos em dois clusters Oracle RAC físicos (4 nós ao todo).
	Todos os appliances existentes e o servidor de backup serão físicos. Todos os demais servidores serão virtualizados: • Microsoft SQL Server • Microsoft Active Diretory • Microsoft Internet Information Services • Apache / Tomcat • Oracle Application Server • Squid Proxy Server • Servidor de arquivos • VMware vCenter Server

E05

A **Tabela 10-12** ilustra o requisito R05 e a especificação E05.

Tabela 10-12 Requisito e Especificação 05

Requisito:	R05
	O sistema de cópia de segurança deve ser confiável, rápido e abranger todos os dados do DATACENTER.
Especificação:	E05
	Utilizar backup em disco para alcançar a confiabilidade e a rapidez requeridas para a política de backup e recuperação.
	Utilizar backup em fita para dados inativos (arquivamento), históricos, backups antigos e política de proteção off-site.

	Fazer backup de: • Arquivos de dados. • Aplicações (Banco de dados Oracle e SQL Server). • Máquina virtual completa (Virtual Disk (vmdk) e demais arquivos).
	Sistema de backup centralizado implantado em servidor físico para fazer backup de servidores físicos e virtuais.
	Solução de backup para cada tipo de dado: • Arquivos – Backup via LAN (agentes instalados na VM ou host físico). • Aplicações – Backup via LAN (agentes instalados na VM ou host físico). • Máquina virtual – Backup de imagem consistente via SAN através do VMware vSphere Data Recovery.

E06

A **Tabela 10-13** ilustra o requisito R06 e a especificação E06.

Tabela 10-13 Requisito e Especificação 06

Requisito:	R06
	As proteções operacionais e de desastres devem permitir recuperação contra falhas lógicas sem perda de dados, ou seja, permitir flashback dos dados.
Especificação:	E06
	Utilizar sistema de armazenamento de journal por operação de I/O para permitir a criação de discos (device, partições ou LUNs) de uma imagem anterior à falha física ou lógica, através da especificação do dia e horário da imagem. Utilizar o sistema de continuidade de negócio e replicação remota EMC RecoverPoint integrado ao EMC Clariion CX4 e compatível com todas as aplicações do ambiente.

E07

A **Tabela 10-14** ilustra o requisito R07 e a especificação E07.

Tabela 10-14 Requisito e Especificação 07

Requisito:	R07
	O novo DATACENTER deve atender aos níveis de serviço (SLA) de até 99,9%.
Especificação:	E07
	Sistema de armazenamento alto disponível para armazenamento dos servidores virtuais.
	Sistema de alta disponibilidade contra falhas de servidores físicos e lógicos.
	Sistema de recuperação de desastre com a utilização de DATACENTERS redundantes distanciados em 5 km.

E08

A **Tabela 10-15** ilustra o requisito R08 e a especificação E08.

Tabela 10-15 Requisito e Especificação 08

Requisito:	R08
	Repositório global de todos os DATACENTERS de arquivos NFS para o ambiente virtual e CIFS para os repositórios de arquivos dos usuários.
Especificação:	E08
	Sistema de armazenamento completo que permite entregar discos de bloco Fibre Channel e iSCSI, além de dispositivos de armazenamento em rede NFS, CIFS e FTP com o EMC Celerra NS.

A **Tabela 10-16** ilustra o requisito R09 e a especificação E09.

E09

Tabela 10-16 Requisito e Especificação 09

Requisito:	R09
	Facilidade na atualização do banco de dados de homologação e teste. OBS: No processo atual é efetuado backup e recuperação do banco de origem para o destino através dos utilitários export/import do Oracle.
Específicação:	E09
	Utilizar recursos de cópia completa ou de ponteiros (snapshot) do sistema de armazenamento. Com este recurso, a atualização do banco de dados será feita de forma mais rápida e fácil, sem a necessidade de recursos de backup/host.

10.7. Dimensionamento da Infraestrutura

Para o dimensionamento da infraestrutura foram utilizados os dados fornecidos pelo cliente e levantados no diagnóstico e foram utilizadas ferramentas de sizing comentadas no capítulo 9 e experiência já adquirida com este tipo de projeto.

10.7.1. Dimensionamento dos Servidores

A **Tabela 10-17** ilustra a taxa de consolidação e a plataforma especificada.

Tabela 10-17 Dimensionamento da Plataforma e Taxa de Consolidação

Plataforma especificada	Taxa de consolidação
Servidor Blade Dois sockets, quadcore Xeon 5530 2,4 Ghz 32 GB RAM Quatro NICs Gigabit Ethernet Duas HBA 4 Gb Dois discos SAS 10K RPM	10:1

- Utilização total de 4 lâminas de blade participantes de um único cluster VMware.
- Emprego do Hypervisor ESXi com boot via disco local.

10.7.2. Dimensionamento do Armazenamento

A **Tabela 10-18** ilustra o dimensionamento do armazenamento.

Tabela 10-18 Dimensionamento do Armazenamento

Área	Utilização	RAID	Desempenho	Espaço líquido
01	Oracle RAC #1	R10	> 1.200 IOPS	1.200 GB
02	Oracle RAC #2	R10	> 1.200 IOPS	1.200 GB
03	NAS (NFS, CIFS)	R5	~ 700 IOPS	900 GB
04	Ambiente Virtual (VMFS e RDM) – Objetos replicados e não replicados	R5	> 1.200 IOPS	2.400 GB
05	Reservado para política de proteção (snapshots e journal/logging)	R5	> 1.200 IOPS	2.400 GB

10.7.3. Dimensionamento da Alta Disponibilidade

A **Tabela 10-19** ilustra o dimensionamento da alta disponibilidade.

Tabela 10-19 Alta Disponibilidade

Tipo de falha	Solução projetada	Funcionalidade
Inoperância do host – Desligamento ou inatividade total do host devido a falha catastrófica de algum componente físico principal ou falta de alimentação elétrica.	Cluster Failover	VMware HA
	Cluster Zero Downtime (Nível 1)	VMware FT
Inoperância de VM – Congelamento ou desligamento do sistema operacional por falha inesperada, paralisando o serviço.	Cluster Failover de VMs	VMware HA
Falha de canal de rede LAN – Falha de qualquer componente em um canal de rede como, por exemplo NIC, switch ou cabeamento.	Alta disponibilidade de caminhos	Switch Virtual
Falha de canal de rede SAN – Falha de qualquer componente em um canal de rede como, por exemplo, HBA, switch ou storage processor.	Alta disponibilidade de caminhos com duas SANs	Failover nativo do VMware

Falha em sistema de armazenamento	Redundância de todos os componentes (fontes, SPs, discos, etc)	EMC Celerra NS-120
Desastre no DATACENTER principal	Site de contingência	EMC RecoverPoint e Celerra Replicator

10.7.4. Necessidade do Uso de Recursos

A **Tabela 10-20** ilustra a necessidade de uso dos recursos.

Tabela 10-20 Necessidade de Uso dos Recursos

Tipo de gestão	Solução projetada	Funcionalidade
Garantia de recursos de processamento e memória	Reserva de CPU e RAM	Pool de recursos VMware
	Balanceamento	VMware DRS
Garantia de recursos de armazenamento	Reserva de IO, banda e tempo de resposta	Navisphere Quality of Service
Redução de custos de energia	Consolidação automática de VMs e desligamento de hosts	VMware DPM

10.7.5. Desenho do Armazenamento

A **Tabela 10-21** ilustra o desenho do armazenamento.

Tabela 10-21 Desenho do Armazenamento

Necessidades	Solução
Conectividade Fibre Channel	Acesso aos arquivos da máquina virtual armazenados em Datastores VMFS e RDM.
Conectividade iSCSI	Acesso a LUNs configuradas dentro de máquinas virtuais ou acesso de hosts físicos para demais ambientes.
Conectividade NFS	Armazenamento de Templates e ISOs acessíveis a todos os DATACENTERS da instituição.

Conectividade CIFS	Servidor de arquivos virtualizado dentro do storage. Elimina necessidade de recursos de host e gerência de máquina virtual. Aumenta flexibilidade e recursos para o serviço de servidor de arquivos.
Datastores	• Template_ISO (NFS) • Nao_replicado (swap e temporários) (VMFS) • Produção_01 (VMFS) • Produção_02 (VMFS) • Produção_03 (VMFS) • Desenvolvimento_01 (VMFS)
RDM	Utilizado nos dados de aplicações que necessita de gerência mais granular no storage e/ou cluster virtual físico.

10.7.6. Dimensionamento da DMZ

A **Tabela 10-22** ilustra a solução de DMZ.

Tabela 10-22 Segurança da Informação

Necessidades	Solução
Isolamento dos servidores interno e DMZ	vSphere Distributed Virtual Switch com PV-LAN

10.7.7. Desenho do Gerenciamento e Monitoramento

A **Tabela 10-23** ilustra o desenho do gerenciamento e monitoramento.

Tabela 10-23 Desenho do Gerenciamento e Monitoramento

Necessidades	Solução
Tipo do vCenter	Virtualizado no primeiro host ESX
Banco de dados vCenter	Banco de dados Oracle 10g no ambiente RAC
Números de vCenter	Um sistema para cada DATACENTER (principal e contingência)
Usuários e grupos	Integrado ao Active Directory

Controle de permissões	Dividir cluster e direcionar VMs: • Pasta produção – Administrador geral do DATACENTER • Pasta desenvolvimento – Administrador desenvolvimento
Administrador de produção	Pertencente ao grupo administrador do DATACENTER
Administrador de desenvolvimento	• Pasta desenvolvimento – Acesso total. • Datastore desenvolvimento – Acesso total. • Datastore template_ISO – Leitura.
Alarmes	• Host storage status. • Datastore usage on disk. • Host hardware fan status. • Host CPU usage. • Network connectivity lost. • Host hardware temperature status. • Host memory usage. • Network uplink redundancy lost.

10.8. Definição dos Níveis de Serviço e Indicadores

As tabelas a seguir detalham os níveis de serviço e indicadores do projeto para servidores, unidades de armazenamento, recuperação operacional e recuperação a desastres.

Tabela 10-24 Servidores

	ATRIBUTOS DE ALINHAMENTO		NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4
Servidores	Desempenho	% de processador garantido	100%	50%	33%	
		% de RAM física garantida	95%	85%	75%	
	Disponibilidade	Tempo máximo de inatividade (ano)	< 525 min	< 4300 min	< 4300 hor	

Tabela 10-25 Storage

		ATRIBUTOS DE ALINHAMENTO	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4
Armazenamento principal	Desempenho	Throughput de desempenho (IOPS)	1.200+	Até 900	Até 500	Até 500
		Tempo de resposta (ms)	< 8	7 a 14	15 a 30	15 a 30
	Disponibilidade	Tempo máximo de inatividade não planejada por ano (minutos)	< 30	< 30	< 60 n	< 240

Tabela 10-26 Recuperação Operacional

		ATRIBUTOS DE ALINHAMENTO	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4
Recuperação operacional	Característica	Classificação da recuperação	Restauração completa do aplicativo	Restauração completa do aplicativo	Restauração de arquivo ou sistema de arquivos	Restauração de arquivo ou sistema de arquivos
	RPO Operacional	Volume de perda de dados	0 min	6 hor	24 hor	5 dias
	RTO Operacional	Tempo necessário para recuperação	< 30 m	< 90 m	7 GB/m	0,5 GB/m
	Capacidade de recuperação	Capacidade de recuperar dados de backup	100%	100%	98%	95%
	Período de retenção	Tempo de retenção dos dados	4 Sem	4 Sem	2 Sem	10 DIAS

Tabela 10-27 Recuperação a Desastres

		ATRIBUTOS DE ALINHAMENTO	NÍVEL 1	NÍVEL 2	NÍVEL 3	NÍVEL 4
Recuperação de desastres	RPO Desastre	Volume de perda de dados	5 min	< 4 hor	24 hor	24 a 48 hor
	RTO Desastre	Tempo de restauração dos dados	< 2 hor	< 12 hor	< 48 hor	< 72 hor

As fases de implementação que envolvem a aquisição de produtos e serviços, instalação e testes da solução não são objetivo deste capítulo.

10.8.1. Diagramas

As **Figuras 10-2** e **Figura 10-3** ilustram os DATACENTERS principal e remoto de forma conceitual.

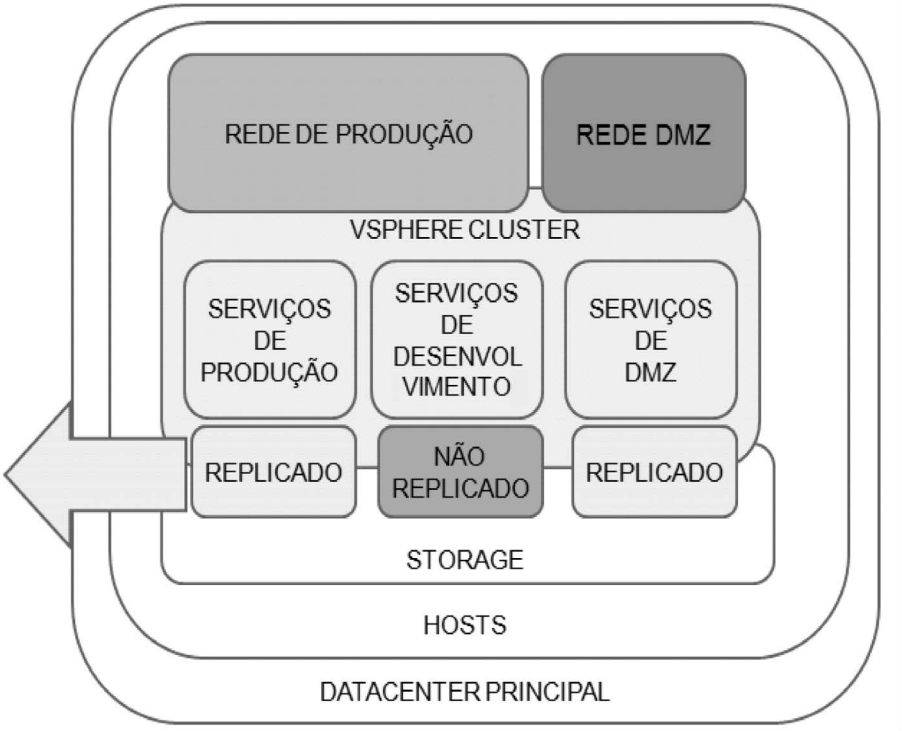


Figura 10-2 DATACENTER principal

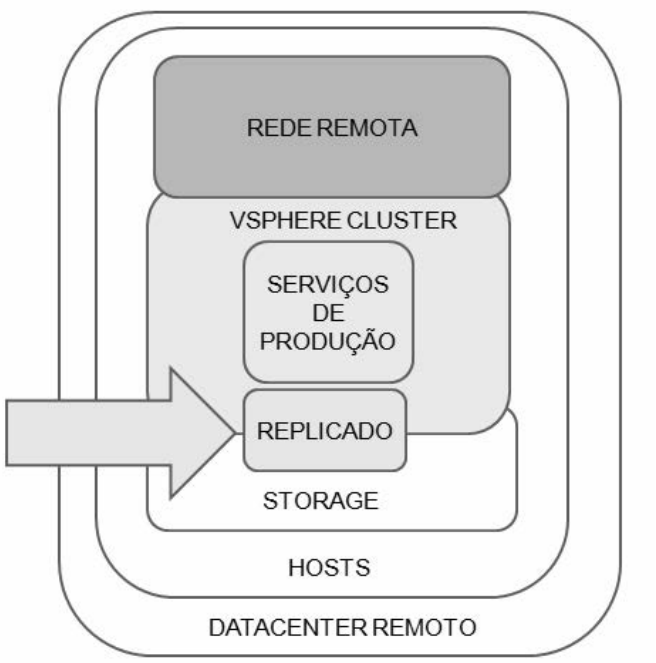


Figura 10-3 DATACENTER remoto

As **Figuras 10-4**, **Figura 10-5** e **Figura 10-6** ilustram os diagramas lógicos dos sites principal e remoto e o diagrama físico do DATACENTER.

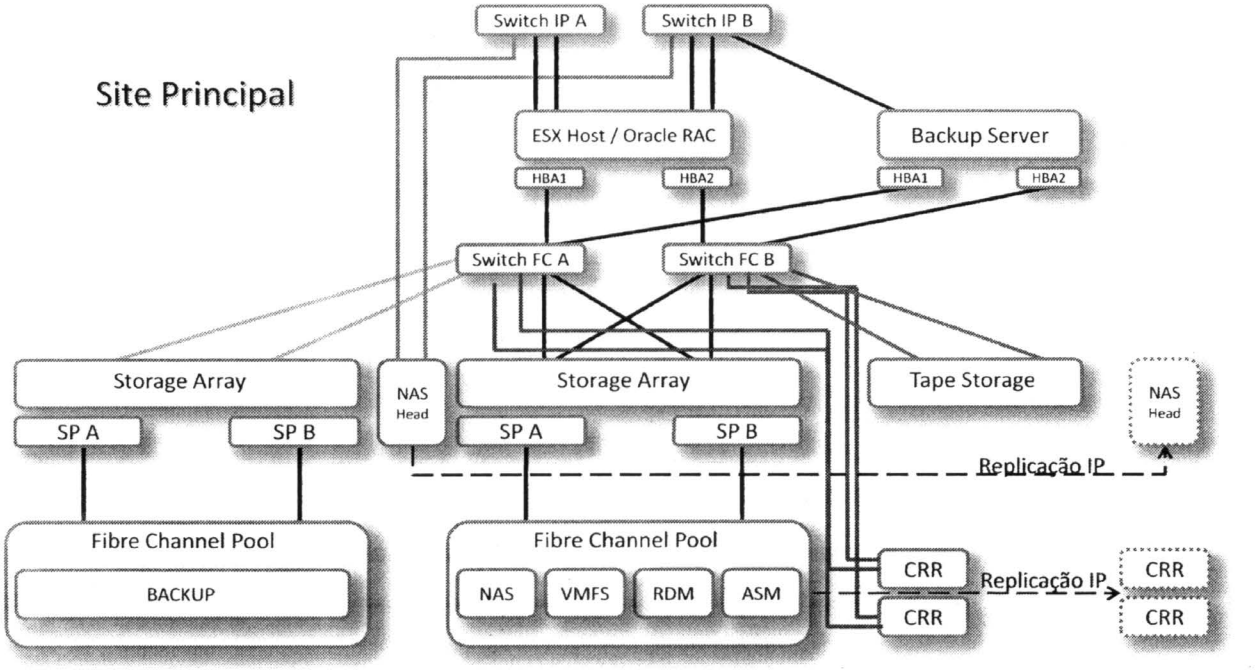


Figura 10-4 Diagrama Lógico do DATACENTER principal

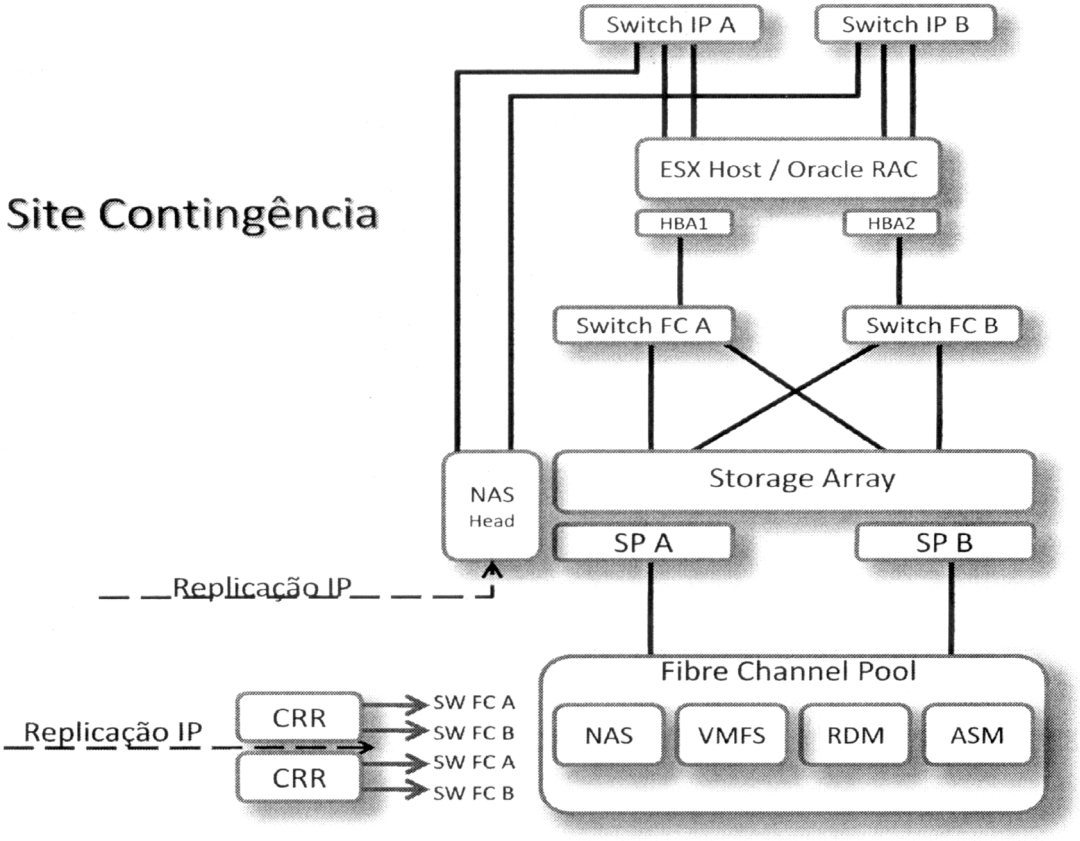


Figura 10-5 Diagrama Lógico do DATACENTER de Contingência

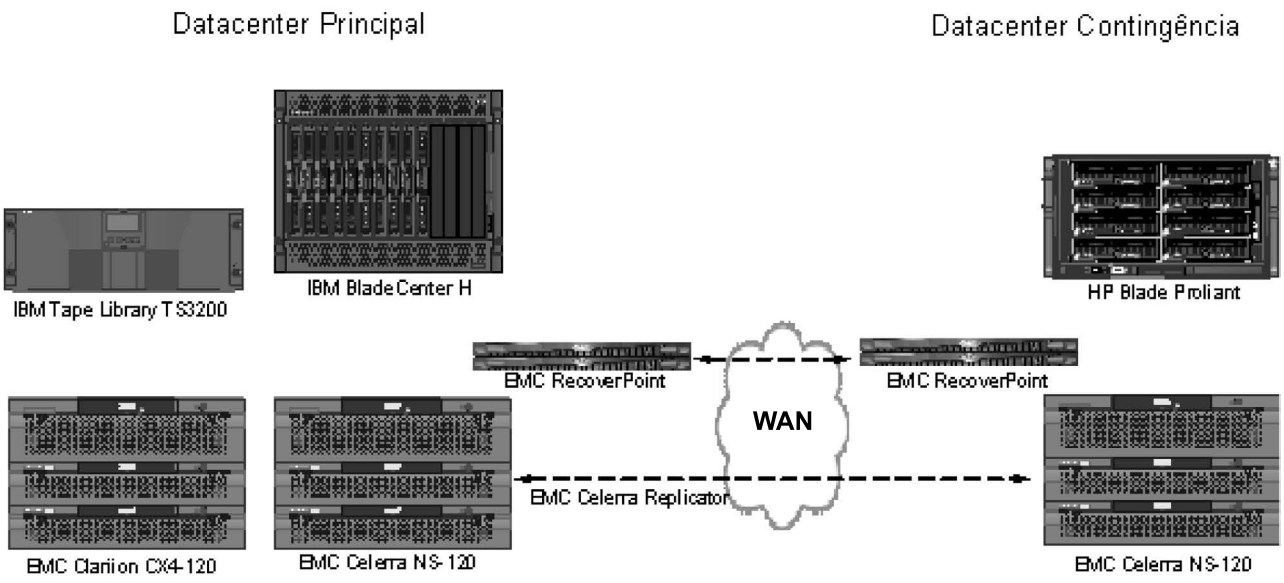


Figura 10-6 Diagrama Físico

10.9. Referências Bibliográficas

Building the Virtualized Enterprise, VMware white paper, 2006.

Capacity Planning: The Blueprint of Server Consolidation, VMware, 2006.

Reducing Server Total Cost of Ownership with VMware Virtualization Software, VMware, 2006.

The True Cost of Virtual Server Solutions, Taneja Group, 2009.

Veras, Manoel. DATACENTER: Componente Central da Infraestrutura de TI, Brasport, 2009.

Virtualization Reality: Why Microsoft Virtualization Solutions Deliver Value when Compared to VMware, Microsoft, 2009.

Why Choose VMware? white paper, VMware, 2010.

Índice Remissivo

Símbolos

64 bits 125, 276, 282

A

ACECO TI 64
AD 15, 177, 284
AMD 71, 125, 182, 265
aplicações 1, 7, 9, 14, 22, 25, 27, 30, 43, 49, 50, 51, 52, 53, 54, 87, 88, 103, 104, 136, 137, 142, 144, 154, 161, 169, 176, 243, 265, 266, 268, 272

B

Backup 165, 167
benchmark 266
Biblioteca de Infraestrutura de Tecnologia da Informação 18

C

capacidade 13, 15, 49, 57, 71, 79, 130, 138, 163, 169, 171, 241, 242, 243, 263, 264, 265, 275, 276, 278, 283, 286
Capacity benchmarking 264
Capacity modeling 264
capacity planning 244, 263, 264, 265
Capacity trending 264
CAPEX 27, 30, 58
CHASSIS 59
CIO 91, 246
Cisco 182
cliente/servidor 126, 137
Cluster 165, 281
CommVault 167
concorrência 13
Consolidação 273, 275
CONTAINERS 61
continuidade 67, 162, 166, 169
custo total de propriedade 21, 23, 50, 171

D

DAS 140, 143, 147, 155
DATACENTER 1, 2, 25, 26, 27, 48, 49, 50, 51, 52, 53, 54, 55, 56, 62, 64, 67, 68, 71, 72, 76, 79, 80, 81, 82, 90, 113, 136, 139, 149, 165, 177, 242, 243, 244, 263, 264, 281, 284, 285, 286
DELL 60, 68, 292
desastres 49, 52, 166, 169, 171, 286, 291
Desempenho 50, 124
disponibilidade 1, 30, 52, 53, 57, 67, 71, 91, 124, 139, 147, 156, 159, 160, 161, 162, 163, 164, 165, 216, 225, 246, 275, 281, 286, 288

E

EMC 23, 66, 90, 182, 226, 234
Escalabilidade 50, 124, 125, 273
estratégia 1, 67, 129, 166, 167, 264, 272

Ethernet 120, 141, 143, 144, 149, 150, 165, 182
ETL 276, 279

F

FCIP 182
Fibre Channel 138, 140, 143, 149, 182
fita 55, 154, 167, 169, 170
Flexibilidade 272

G

Gartner 21, 42, 242
Gerenciamento 23, 65, 120, 124, 273
Green Grid 71, 72, 76, 81, 82

H

HP 58, 66, 182
HPC 149, 164, 266
HYPERVISOR 88, 90, 101, 104, 106, 108, 177

I

IBM 89, 129, 130, 292
IDC 48, 66
iFCP 182
Information Technology Infrastructure Library 18
infraestrutura de TI 1, 2, 7, 8, 9, 11, 13, 14, 17, 18, 21, 22, 26, 67, 113, 130, 136, 165, 263, 264, 289
inovação 139
integridade 172, 286
INTEL 265
interoperabilidade 72, 154
investimento 9, 11, 13, 125, 148, 161, 169
IOM 14
iSCSI 138, 142, 143, 144, 155, 182, 231, 234
ITIL 18, 21, 22, 23
ITIL v2 18
ITIL v3 18

K

KVM 55

L

LINPACK 266
Linux 22, 288
LUN 232, 287

M

Microsoft 14, 165, 267, 272, 275, 276, 284, 292
Modelo de Otimização de Infraestrutura 14
MOLAP 278
MS Exchange 267, 282, 284, 285, 287
MS SQL Server 273, 275, 279, 288
MTBF 53

N

NAS 140, 141, 144, 147, 148, 154
Nexus 149

O

OLAP 276, 278
OLTP 266, 268, 276, 279, 281
OPEX 30, 58, 59
ORACLE 281, 292

P

PaaS 30
Padronização 59
portfólio de TI 11
processos de negócio 25

R

RACKS 55, 61, 64, 71, 242, 243, 244
RAID 124, 139, 140, 171, 231, 275
Real Application Clusters 281

Replicação 171, 172, 173
RISC 116, 125, 265
risco 15, 42, 263, 288
ROLAP 278

S

sala cofre 63, 64
SAN 52, 55, 56, 140, 141, 142, 143, 144, 149, 154, 155, 156, 177, 280, 290
SAP 288, 289, 290, 291, 292
Server 2008 165, 272, 273, 274, 276, 277, 278, 279
sizing 279, 287, 288, 291, 292
SLA 53, 170, 288
SPC 268
SPEC 124, 126, 268
storage 50, 52, 55, 57, 58, 136, 137, 138, 139, 140, 141, 142, 143, 144, 154, 155, 165, 167, 168, 170, 171, 231, 232, 243, 275, 279, 280, 281, 282, 287, 293
Sun 66, 242

T

TCO 21, 22, 23, 50, 53, 58, 59, 67, 71, 90, 144, 154, 176, 242, 271
TCP/IP 141, 143, 148
The Green Grid 71, 81, 82
TIA 942 54, 55, 57
TIER 57, 58
Total Cost of Ownership 21, 66
TPC 124, 126, 266, 267
TPC-C 126
TPC-E 126, 266

U

UNIX 148
Uptime Institute 53, 54, 66, 67

V

virtualização 1, 51, 52, 80, 154, 155, 156, 177, 274, 292
VM 88, 90, 114, 130, 145, 213, 225, 266
VMware 34, 71, 93, 101, 129, 130, 145, 177, 185, 216, 223, 224, 225, 226, 228, 231, 232, 234, 235, 262, 269, 271, 295
vSphere 34

W

Windows 22, 129, 148, 161, 165, 176, 273, 284
Windows Server 2003 273

X

x86 90, 91, 92, 108, 116, 117, 124, 125, 126, 129, 130, 176, 247, 265, 281